

Kentucky Department of Education Abstract

Gender and Race Subgroup Performance Difference in KIRIS Accountability Cycle 2 and Cycle 3 Results

Kentucky's statewide assessment and accountability program, the Kentucky Instructional Results System (KIRIS), was designed to foster learning for all children in the Commonwealth. A recent KDE Office of Assessment and Accountability study examined key parts of the KIRIS assessment to explore the extent to which there were performance differences by gender and race. Data from all six years of KIRIS testing were included in the analyses (1992-98). The main findings of this study were that small academic performance differences by gender and race exist among Kentucky students. The overall pattern of difference found the following rank ordering of student scores (highest to lowest): Caucasian female, Caucasian male, African-American female and African-American male. The following paragraphs present brief summaries of the results by school level followed a brief examination of the performance differences.

Elementary Schools

The overall elementary KIRIS scores indicated continuous improvement across the six-year period in reading, mathematics, science and social studies scores. There was very little difference by gender in the science and mathematics scores; however, there were small differences by gender in the reading and social studies scores. The differences in scores by race were small but consistent in all academic areas across all six years.

Middle School

The overall middle-school KIRIS scores for the six-year period indicated a steady state of performance in reading, science, and social studies and a slight increase in mathematics. This pattern was consistent across groups, but there were small differences between group performances. There were differences by both race and gender in the reading and social studies scores; however, the group differences in mathematics and science scores were by race rather than gender. Caucasian females and Caucasian males had very similar science and mathematics scores; these scores were higher than the scores of African-American females and African-American males.

High Schools

The high school KIRIS scores indicated progress in academic achievement across the six-year period in all areas. The pattern of group difference in the high school science and mathematics scores was similar to that of middle-school scores. The difference was by race rather than gender. There were small group differences in the high school reading and social studies scores that followed the overall pattern with the groups rank ordered as Caucasian female, Caucasian male, African-American female, and African-American male.

Examining Performance Differences

Comparing patterns of responses with national assessments.

Examination of the available data from the National Assessment of Educational Progress (NAEP) and the ACT revealed a similar pattern of performance differences between males and females and between Caucasian and African-American students. Education Week reported that the race difference decreased from 1970 to 1988 and has been increasing ever since (Olson, 1997). The small differences in KIRIS scores were consistent across the six-year period; indicating neither increases nor decreases in differences.

Examining the test for bias.

Although there are performance differences evident within the KIRIS data between males and females and between Caucasians and African-Americans, the size of the performance differences and the small amount of variance they explain do not necessarily imply that the KIRIS assessment or the accountability system is biased. The Bias Review Committee reviews the KIRIS assessment extensively during the test construction each year. Items judged by the Bias Review Committee to cause disadvantage to student subgroups are dropped from the test or modified to remove the potential source of bias.

Differences in performance by race in districts with high concentrations of minorities.

An additional section of the study compared the gap in Caucasian and African-American students' scores in school districts with higher concentrations of minorities versus districts with lower concentrations of minorities. Results of this preliminary study found that school districts with more than 3% minority enrollment had a larger discrepancy between Caucasian and African-American scores.

Qualitative exploration of racial differences in performance.

An exploratory study examined six Kentucky schools looking for common factors that could explain the small racial performance difference at each school over the Accountability Cycles. Semi-structured interviews were conducted at each sample school. An analysis of the data from all six schools found the faculties of every school commonly reported only one variable that they believed affected scores; socioeconomic status. There was a widespread perception among staff and faculty in the schools under study that when students' social class was similar the performance differences between Caucasian and African-American students either became smaller or disappeared. This finding from the qualitative study was followed-up quantitative analysis of the data; using a multiple linear regression to explore the effects of gender, race, and socioeconomic factors on KIRIS scores. This analysis found that the small differences by race and gender were still apparent after socioeconomic differences were accounted for.

This KDE Abstract is followed by the study by Doug Smith et al. The Doug Smith study begins with the Executive Summary on page 1.

EXECUTIVE SUMMARY

The purpose of this research was to examine key parts of the KIRIS assessment to explore the extent to which there were performance differences by gender and race. The main findings of this study were that small academic performance differences by gender and race exist among Kentucky students. Specifically, women are somewhat more likely to have higher subject area scores than men in almost every subject area examined and at every grade level. These performance differences have decreased slightly and do not appear to be affected by geographic concentration. In addition, whites are more likely than blacks to have higher performance area scores. This gap has widened slightly over the last six years in most subject areas examined¹. The concentration of African-Americans may affect this difference also in certain parts of the state. Other preliminary findings indicate that educators in a small sample of Kentucky's schools believed the racial performance differences were a result of students' social class; however, further quantitative analysis in the report tends to call that assumption into question. Reflection upon the currently available data and quality of indicators of social class and other socio-demographics used in this report to measure performance differences leads to a call for more accurate, valid and reliable socio-demographic measures to be collected at the individual, school and county level. Moreover, these additional and refined socio-demographic measures need to be used in a series of cross-validating qualitative and quantitative studies to develop more reliable, valid, and generalizable information concerning performance differences. Last, some centralized research and/or data storage facility is proposed to facilitate research and decision-making in the Commonwealth.

¹ This statement refers to the data in Table 6 page 27 (footnote added by KDE).

INTRODUCTION

Implemented between 1991 and 1998, Kentucky's statewide assessment and accountability program, the Kentucky Instructional Results Information System (KIRIS), was designed to foster learning for all children in the Commonwealth. As such, it was important that the KIRIS assessment and the educational system that it was designed to examine be fair to all students. One crucial fairness question in any testing program is whether the assessment adversely affects a group or groups of students compared with others. While the Kentucky Department of Education (KDE) takes steps to prevent adverse testing effects during test construction, researchers must also examine test results for signs of differential performance that may indicate adverse testing effects.

If these analyses show that differential performances exist based on shared group characteristics, additional investigation is warranted. Although the student assessment results may be examined for differential performance based on any shared social or demographic characteristics, researchers typically compare student performance between the sexes and among ethnic or racial groups. This report examines the extent of student performance differences in KIRIS results over the last two accountability cycles by sex and race.

Description of KIRIS

At its heart the KIRIS assessment was designed to hold schools accountable. To this end, KDE assessed student learning in many subject areas over several grades using several testing methods. Once processed, the student assessment data were aggregated for the schools and districts for accountability purposes. The accountability system was designed as a series of overlapping four-year cycles.² In Accountability Cycle 2 the combined results for academic years KDE compared the combined results of the 1992-93 and 1993-94 academic years with the combined results of the 1994-95 and 1995-96 (KDE 1997). In turn, Accountability Cycle 3 used the combined results of the 1994-95 and 1995-96 academic years as its baseline, which KDE then compared with the combined data for the 1996-97 and 1997-98 academic years.

KIRIS was a work in progress. There were several changes to the assessment program over the two accountability cycles. Among the most visible changes was the change in the grades that KDE tested. During the first two years of Accountability Cycle 2, KDE assessed 4th, 8th, and 12th grade students in all subject areas, using constructed response questions and other testing methods (e.g., portfolios and on-demand writing prompts). In response to complaints about possible 'senior apathy' KDE moved the 12th grade assessment to 11th grade beginning in 1994-95. Furthermore, concerns about the length of the constructed response testing process and its effects on students' scores at the lower grades led KDE to split the subject areas assessed by constructed response tests across grade pairs starting in 1996-97. From that point on, KDE assessed reading and science in grades 4, 7, and 11 while mathematics and social studies were assessed in grades 5, 8, and 11. Each time that assessment grades changed, KDE equated the

²For technical information on the KIRIS assessment please refer to the KIRIS Accountability Cycle 1 Technical Manual (KDE 1995) and the KIRIS Accountability Cycle 2 Technical Manual (KDE 1997).

student scores on the constructed response tests to make them comparable. While several other areas were assessed through the KIRIS constructed response tests (e.g., arts and humanities), this study examines performance differences by sex or race only in the four major subject areas: reading, mathematics, science, and social studies.

The Gender and Racial Composition of Kentucky Schools

Before examining student test performance, at least a cursory look at the demographics of the students tested over Accountability Cycles 2 and 3 is important. The distribution of students by gender in Kentucky's school population was extremely stable over this six-year period. Data in Table 1, indicates that between 1993 and 1998 males were nearly 52 percent of the population in the lower four grades and about 50 percent of the student population in the 11th grade. The decrease in males is may be a differential dropout rates between males and females. Research done at the national level suggests that males are more likely to drop out of school over time than are females (Calabrese and Poe 1991; NCES 1998). In addition, other research at the national level suggests that some in some geographic areas, a small portion of the difference may be due to higher rates of mortality among adolescent males (Retherford 1975).³

While the distribution of male and female students in the five grades assessed is relatively equal, the racial/ethnic mix is not. Mirroring the Commonwealth's overall racial mix, African-American students (constituting about 9%) and whites (constituting nearly 89%) are the two major racial groups in the accountability grades. All other ethnic/racial groups comprise approximately 2% of the students tested. Given the potential inaccuracy of statistics produced from the small numbers of Hispanics, Asians, and students with other ethnic backgrounds, this examination of differential racial performance will address only differences between white and African-American students.

Measures of Student Performance

The KIRIS assessment yields two student-level scores that can be used to assess performance differences. One is theta, a measure of ability within Item Response Theory (IRT). Theta is created by modeling the response patterns of the students' test answers using a graded response model (Samejima 1969). Student response patterns for each of the four major constructed response area tests (i.e., reading, mathematics, science, and social studies) are analyzed separately to develop a theta score for each area. A continuous variable usually standardized so that its range runs from -3 to +3, theta is highly correlated with the total raw score on the constructed response tests.

³In Kentucky dropout rates are kept by gender and race only at the school and school district level. Therefore any effects of dropout rate on performance differences cannot be examined at this time. The Kentucky Department of Education has just begun to collect dropout rate by race from local education agencies. However, this information is not currently available.

TABLE 1. Gender Distribution in Accountability Grades from 1993 to 1998⁴

Grade	Gender	Academic Year						
		1992-93	1993-94	1994-95	1995-96	1996-97		
4	Male	50.9	50.0	51.7	51.6	51.2	48.8	
	Female	47.8	46.9	47.7	48.0	48.7	51.2	
	Missing	1.4	3.0	0.5	0.2	0.0	0.0	
	Total	48951	49106	48685	47116	46941	46405	
5	Male	GRADES 5 AND 7 WERE NOT TESTED DURING CYCLE 2.					51.7	51.2
	Female						48.3	48.7
	Missing						0.0	0.0
	Total						46973	47078
7	Male						51.5	52.1
	Female						48.3	47.9
	Missing						0.0	0.1
	Total						49441	49094
8	Male	50.0	50.3	51.0	51.0	51.3	51.0	
	Female	48.4	48.1	48.4	48.6	48.7	48.9	
	Missing	2.0	1.7	0.6	0.3	0.0	0.0	
	Total	49018	51302	51090	50244	49521	48539	
11/12	Male	49.3	50.0	49.5	49.3	49.1	48.8	
	Female	49.9	49.5	47.7	50.5	50.9	50.8	
	Missing	0.9	0.4	2.8	0.2	0.0	0.2	
	Total	35832	36388	41082	41118	40819	40327	

⁴The numbers and the percentages presented in the table and in the text were based on self-reported student responses. These numbers do not include students taking alternate portfolios and, therefore, may differ slightly from figures from other sources. In addition, the percentages presented in the table may not add up to 100 due to rounding.

Table 2. Racial/Ethnic Distribution in Accountability Grades from 1993 to 1996⁵

Grade	Race/ Ethnicity	Academic Year						
		1992-93	1993-94	1994-95	1995-96	1996-97	1997-98	
4	Asian	0.3	0.4	0.4	0.5	0.6	0.5	
	African-American	8.1	8.7	8.2	9.7	9.4	10.1	
	Hispanic	0.2	0.3	0.3	0.4	0.5	0.6	
	White	84.9	86.2	86.5	88.1	86.7	87.2	
	Other	0.3	0.5	0.5	0.6	0.6	0.9	
	Missing	6.0	3.9	4.1	0.7	2.2	0.7	
	Total	48951	49106	48685	47116	46941	46405	
5	Asian	GRADE 5 AND 7 WERE NOT TESTED DURING CYCLE 2					0.5	0.6
	African-American						9.4	9.7
	Hispanic						0.5	0.6
	White						86.5	87.9
	Other						0.7	0.8
	Missing						2.4	0.4
	Total						46973	47078
7	Asian	GRADE 5 AND 7 WERE NOT TESTED DURING CYCLE 2					0.5	0.5
	African-American						9.1	9.1
	Hispanic						0.5	0.5
	White						86.7	88.4
	Other						0.9	0.9
	Missing						2.3	0.5
	Total						49441	49094

⁵The numbers presented in this table and the percentages presented in the text were based on the number of self-reported student responses. These numbers do not include students taking alternate portfolios and therefore may differ slightly from figures from other sources.

Table 2. Racial/Ethnic Distribution in Accountability Grades from 1993 to 1996 (Continued)

8	Asian	0.3	0.4	0.4	0.4	0.5	0.6
	African-American	9.3	8.6	8.5	9.1	8.8	9.3
	Hispanic	0.5	0.4	0.5	0.4	0.5	0.6
	White	86.1	87.6	86.6	88.2	86.4	87.9
	Other	0.4	0.7	0.8	0.9	0.8	1.1
	Missing	3.5	2.3	3.2	0.9	3.4	0.5
	Total	49018	51302	50689	50244	49521	48539
11/12	Asian	0.5	0.7	0.5	0.6	0.5	0.6
	African-American	7.4	7.8	7.7	8.5	8.1	8.2
	Hispanic	0.3	0.3	0.4	0.4	0.4	0.6
	White	88.6	89.4	86.5	88.4	87.2	88.9
	Other	0.4	0.8	1.1	1.1	1.1	1.2
	Missing	2.9	0.9	3.8	1.1	2.7	0.6
	Total	35832	36388	41082	41118	40819	40327

The other measure of student performance that can be created from the KIRIS assessment is the constructed response index (CRI), a measure created by KDE for use in school accountability.⁶ Using cutpoints developed by KDE standard setting committees in 1992, which KDE verified with the KIRIS standards validation study in 1995, KDE divides students' theta scores for the constructed response subtests into four categories, Novice, Apprentice, Proficient, and Distinguished (KDE 1995, 1997). Next, KDE assigns each of these categories a numerical weight of 0, 40, 100, and 140, respectively. These subtest indices are an alternate measure of student performance.

Either theta scales or constructed response indices can be used to examine student performance. The constructed response indices (CRI) are congruent with existing educational thought and policy in the Commonwealth because they comprise most of the academic index that

⁶This measure is sometimes referred to in other KDE publications as the performance level index.

KDE uses to hold schools accountable.⁷ Theta, however, provides a more precise way to examine student performance differences across years because it is a continuous variable. The categorization of students' scores by cutpoints turns the continuous theta scale into the discrete four-category CRI and thus constrains variation between students' scores.

The difference between theta and the CRI is important to understand. This can be illustrated with an example. Suppose two hundred students—100 males and 100 females—are assessed in reading at the start of an accountability cycle. These students' scores are then analyzed and each student is given a theta value. Then it is discovered that a performance difference exists in the data between male and female student thetas such that males, on average, have theta scores that are eight units higher than females. Figure 1 represents the hypothetical standard normal distributions of these students' theta scores. Next suppose that the range of the distributions of both male and female students is so small that all males and females fall in the novice category. In such a case researchers would categorize both male and female students as novices and given a score of zero on the reading CRI. Thus, while an 8-unit difference between the male and female means would be evident on the theta scale, no difference between the means would be detected using the CRI scores at Time 1.

To continue the example, suppose in each new academic year the male and female students' scores show an increase over time and that both groups increase at the same rate. This steady rate of change would mean that the 8-unit difference between the males' and females' theta scores would remain. It would also mean that over time the average male student would move from one category to the next, here from Novice to Apprentice, before the average female student. This case is presented at a hypothetical Time 2 in Figure 2.

Although the underlying performance difference between the male and female student means as measured by theta stayed the same, at Time 2 most of the males would be given a value of 40 on the CRI for the subtest while the all female performances would continue to receive a value of zero. Therefore, at Time 2 the difference between the average score for males and females on the CRI would be 35.6. This perceived performance difference would persist in the CRI scores until the average female theta score increased enough to put females into the same category as the average male. Once female students' scores begin crossing over the cutpoint between novice and apprentice, the perceived difference between males and females as measured by the CRI would again diminish (see Figure 3).

The point is not to persuade the state to quit using the CRI but to emphasize that continuous variables (like theta) are, most often, more accurate statistical measures than four category variables. In the above example, a number of assumptions were made about statistical errors that may come from measures like the CRI. For Kentucky's many of these assumptions do

⁷Other portions of the academic index can include the writing portfolios, mathematics portfolios, the arts and humanities constructed response items, and the practical living and vocational studies constructed response items, depending upon the particular accountability grade and year.

Figure 1. Hypothetical Distribution of Male and Female Student Scores at Time 1

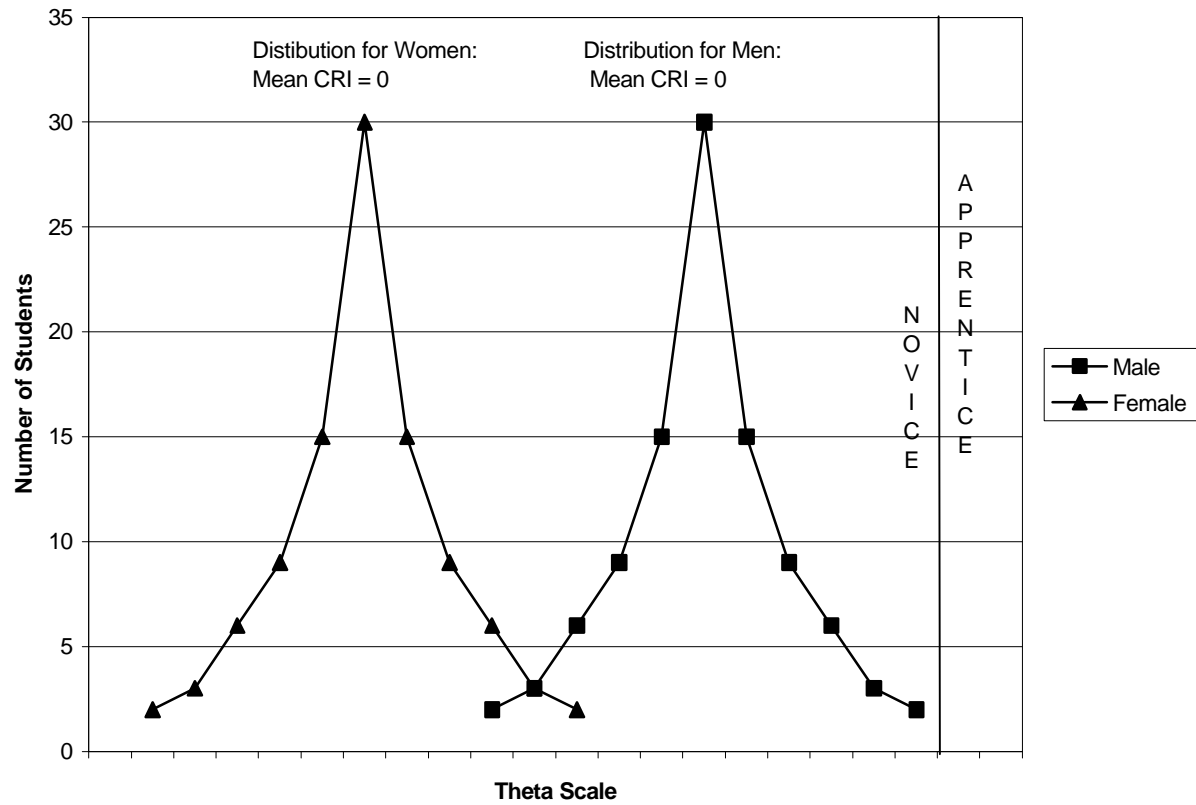


Figure 2. Hypothetical Distribution of Male and Female Student Scores at Time 2

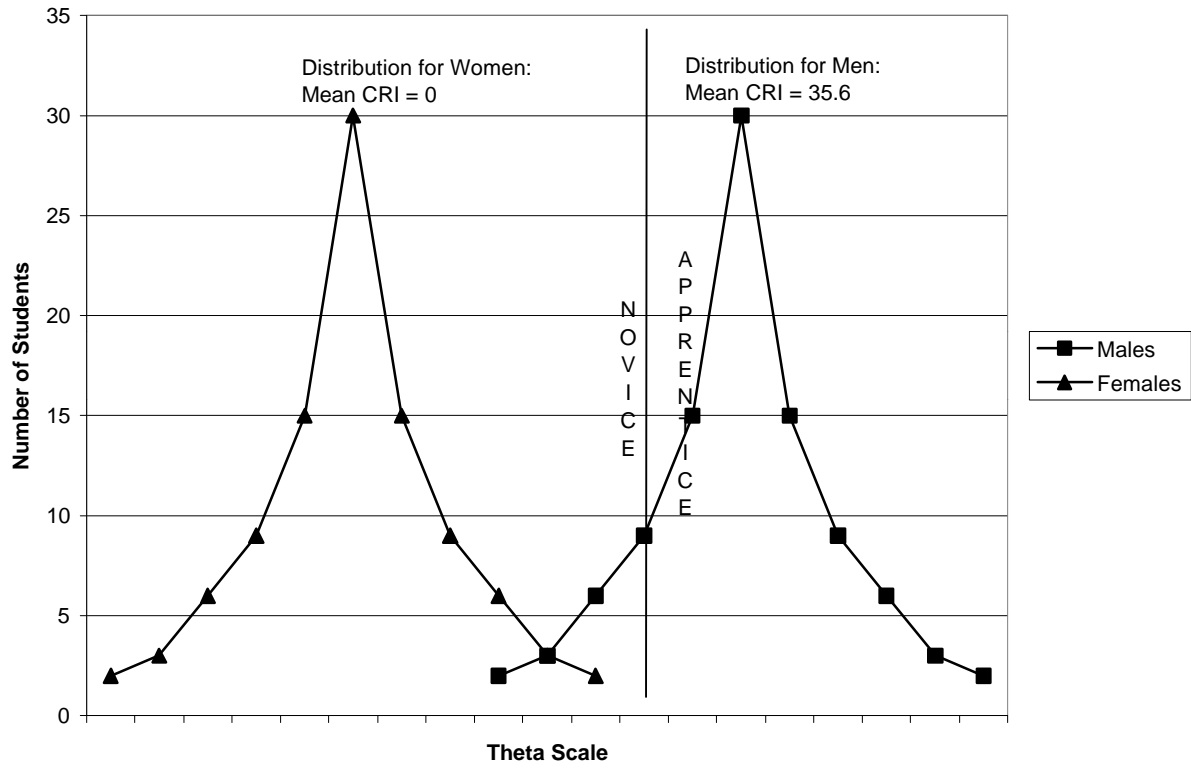
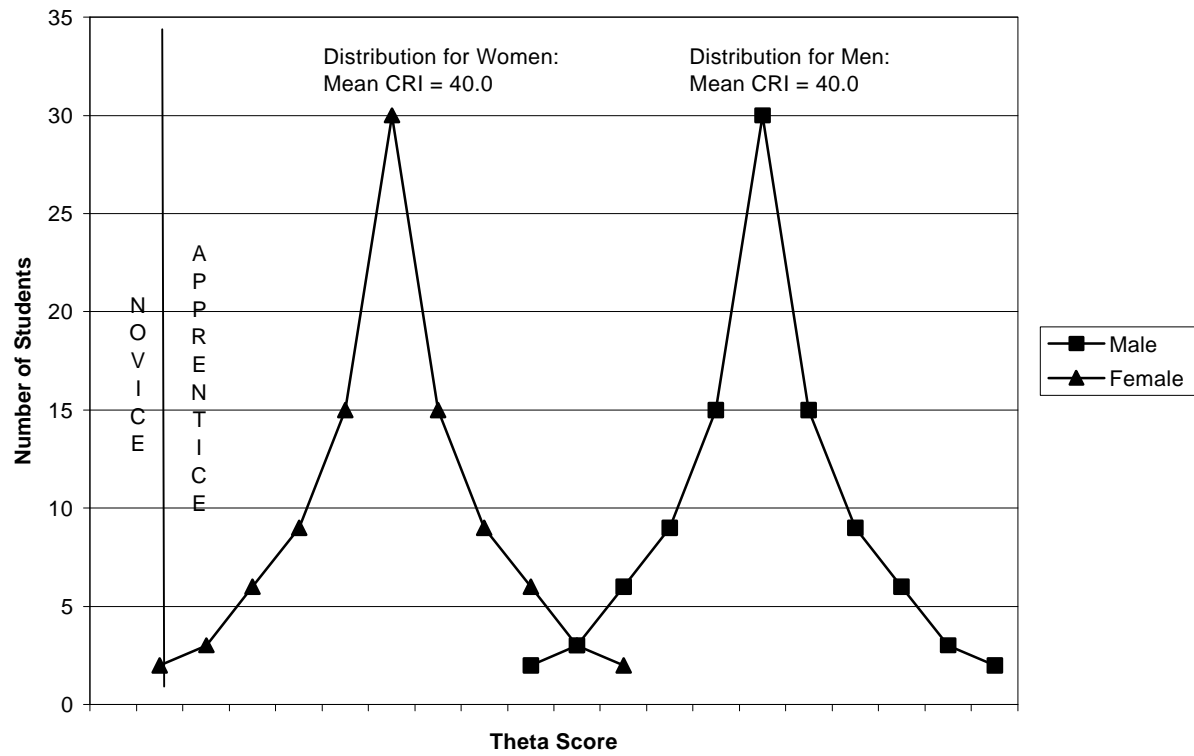


Figure 3. Hypothetical Distribution of Male and Female Student Scores at Time 3



not apply. Kentucky's student population is between 200 and 250 times larger than that used in the example, and Kentucky's students have always been distributed across all four CRI categories in each subtest thus reducing the possible effects of calculation effects in assigning scores to students. An examination of the findings presented in this analysis suggests that there is little difference between the results calculated using theta and those calculated using the CRI scores. Still, while the categorical nature of the CRI is useful from a policy standpoint, the point value awarded for each category of a constructed response index could exaggerate underlying performance differences between the two groups when these groups are in different categories and underestimate these same differences when the groups are in the same categories. Because of such fluctuations, cautious interpretation of the constructed response indices over time is necessary.

While the theta score and the Constructed Response Index score are two measures that can be derived from student data, student performance is often analyzed by comparing the percentage of students scoring in particular CRI categories. Most often this comparison is between subgroups' novice percentages or between their proficient and distinguished percentages. Newspaper articles on race and gender performance differences often display percentages of students scoring proficient and distinguished by race and gender. While such a technique is not necessarily wrong, it can lead to inaccurate interpretations if only part of the distribution is presented for examination. If the shape of the distributions for each group is different, or if shifts in the distribution occur over time but not in the categories being compared, inaccurate conclusions may be drawn. Therefore, when comparing student performance by examining differences in CRI category percentages, presenting the percentages from all four categories is necessary.

EXAMINATION OF STUDENT PERFORMANCE BY GENDER AND RACE

Figures 4 through 15 present reading, math, science, and social studies subgroup performance on the theta scale for the accountability grade assessed by race and gender.⁸ These figures suggest several things about performance on the KIRIS assessment.

First, the analyses indicate that, on the whole, student performance for all four groups has improved over Accountability Cycle 2 and Cycle 3. There are two exceptions to this trend, which to be noted. Figure 7 shows that 7th grade science scores for all four groups have declined. More disconcerting is the data from 8th grade social studies (see Figure 8), which shows that African-American females started even with white males in 8th grade social studies, but their performance has declined over the last three academic years, reaching a new low.

Second, performance differences are evident within the data no matter which subject area is examined. The performance order of these four groups varies little over the last six years. White females have the highest average performance followed by white males and then African-American females. African-American males turn in the lowest average performance in each grade.

Last, there appear to be two distinct performance differences; differential performance cannot solely be laid at the feet of either race or gender. When holding race constant, females, on average, score higher than their male counterparts. This difference is especially evident in reading and social studies. When holding gender constant, the average white performance is higher than the average black performance. This racial difference can be seen in each figure and is most obvious in math and science, where gender differences in scores can be seen to be the smallest.

Performance Differences: Are They Significant?

It is usually the case in social science that the researcher must observe the behavior of a small sample of individuals to infer the behavior of the population of interest. Faced with such a situation, the researcher must spend time making sure that his or her observations are statistically and substantively significant. That is, that the differences observed in the data are actually differences in the population and not differences due to an abnormal sample. Moreover, the researcher also must determine that differences are large enough to affect behavior in real world settings. Faced with these problems, social scientists and applied statisticians have developed the field of inferential statistics. However, because KDE assesses every student

⁸Performance differences measured by theta for subject subtests by sex and race separately can be found in Appendices A and B. Performance differences using CRI scores are in Appendices C, D, and E. CRI Category percentages showing changes in the distributions of males, females, whites, and blacks are found in Appendices F, G, H and I.

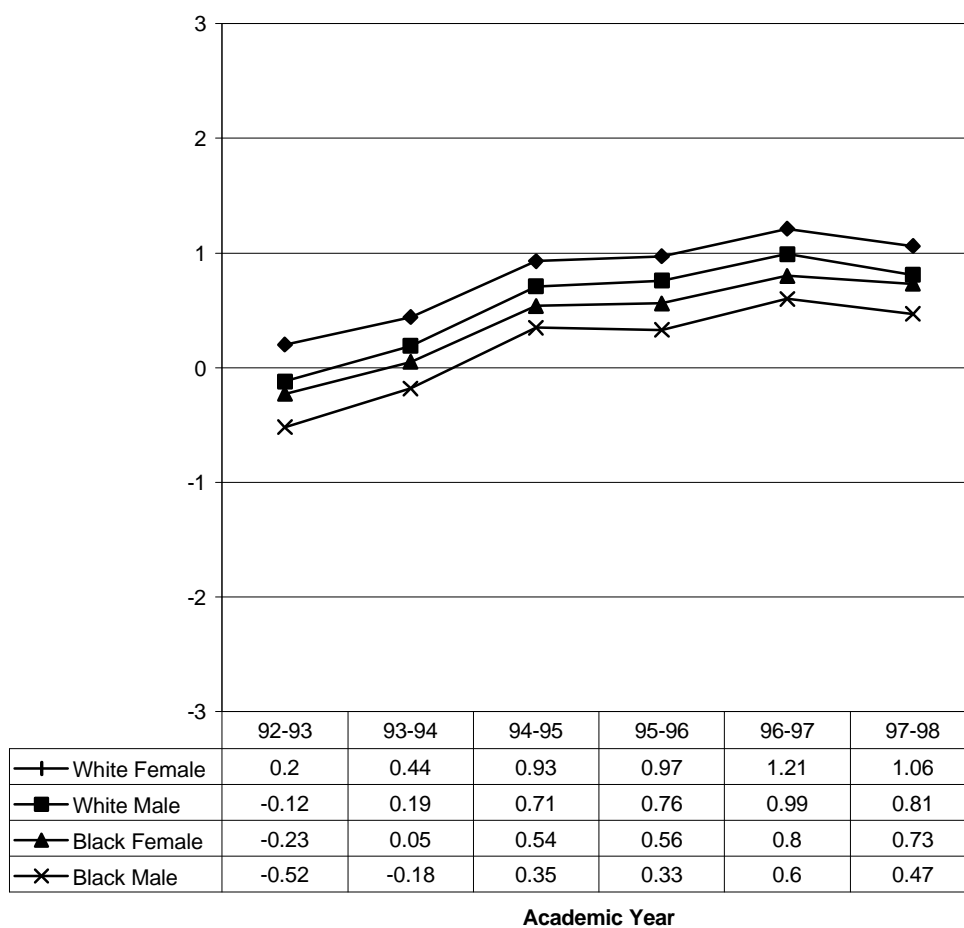
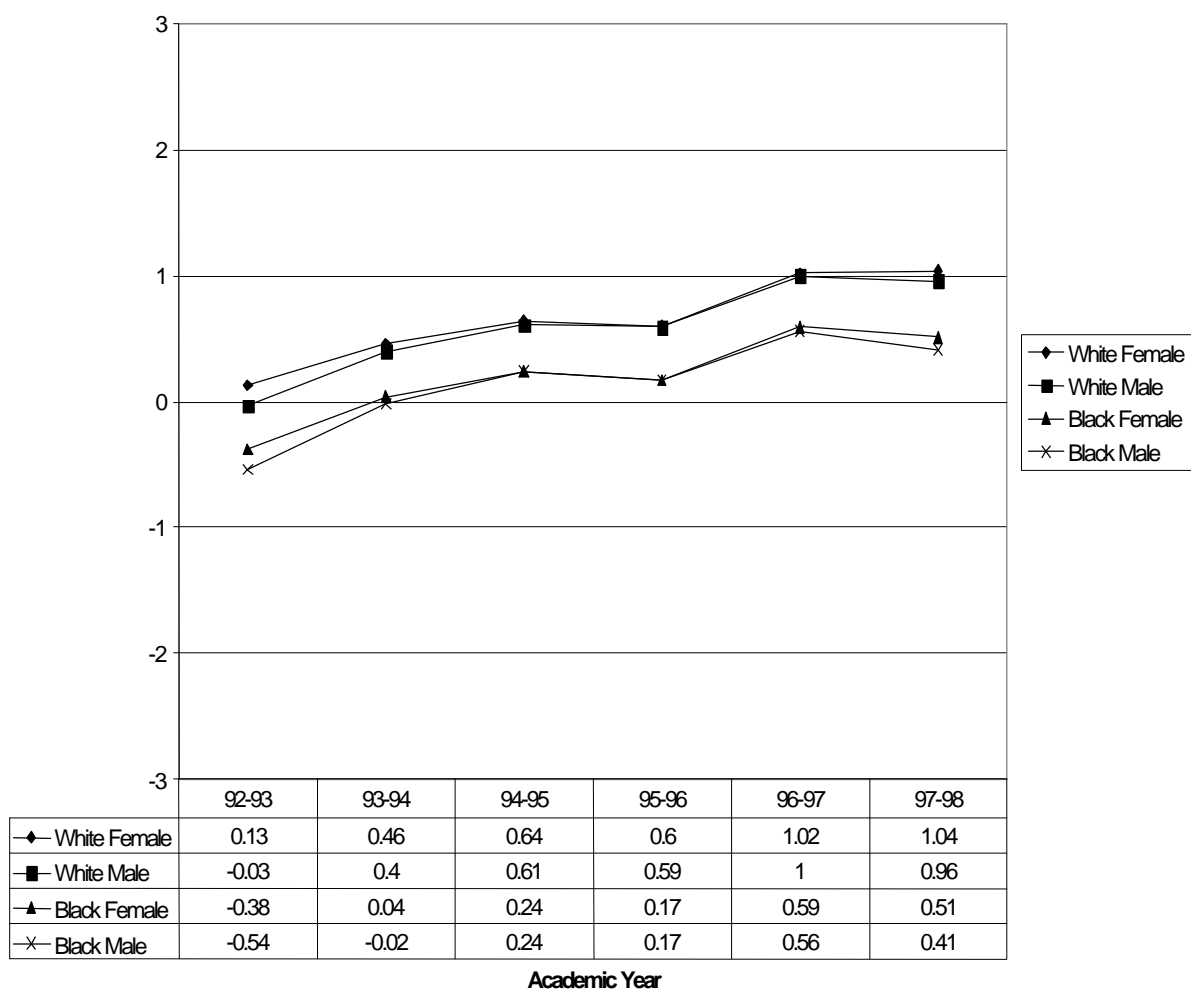
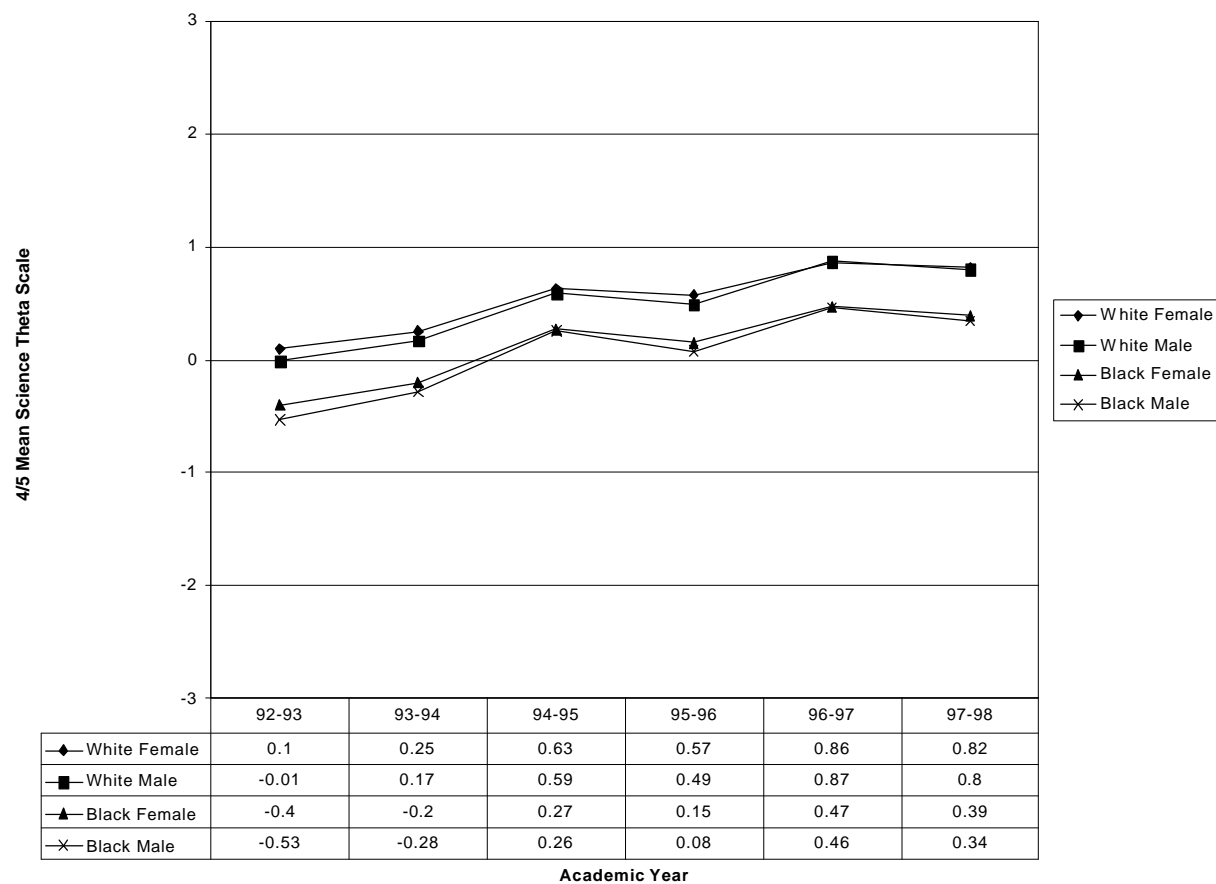
Figure 4. KIRIS Reading Theta for Grade 4 by Gender and Race, 1993-94⁹⁹ Data Table was added to Figure 4 by KDE.

Figure 5. KIRIS Mathematics Theta for Grade 4/5 by Gender and Race, 1993-1998¹⁰

¹⁰ Data Table was added to Figure 5 by KDE.

Figure 6. KIRIS Science Theta for Grade 4 by Gender and Race, 1993-1998¹¹

¹¹ Data Table was added to Figure 6 by KDE.

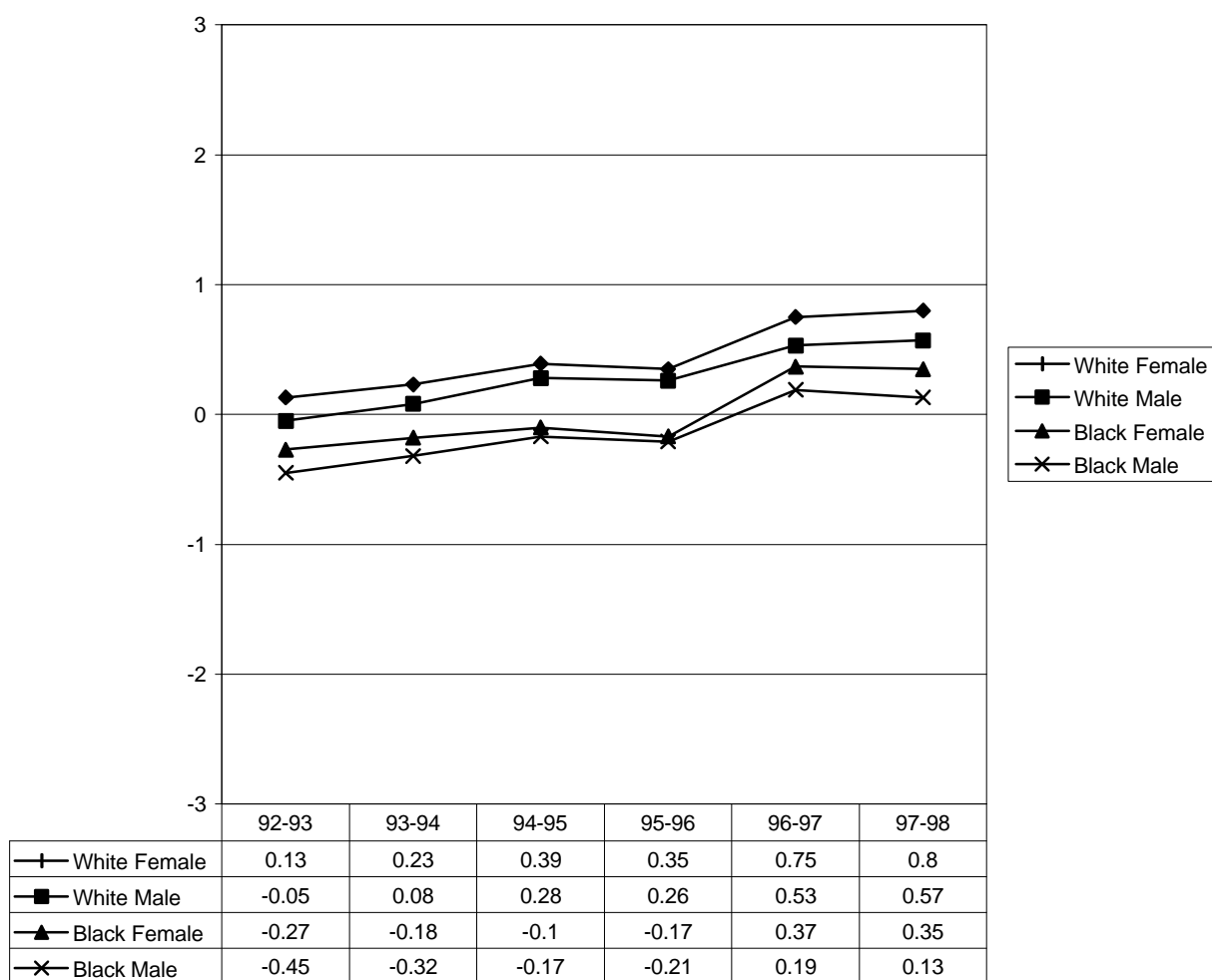
Figure 7. KIRIS Social Studies Theta for Grade 4/5 by Gender and Race, 1993-1998¹²¹²Data Table was added to Figure 7 by KDE.

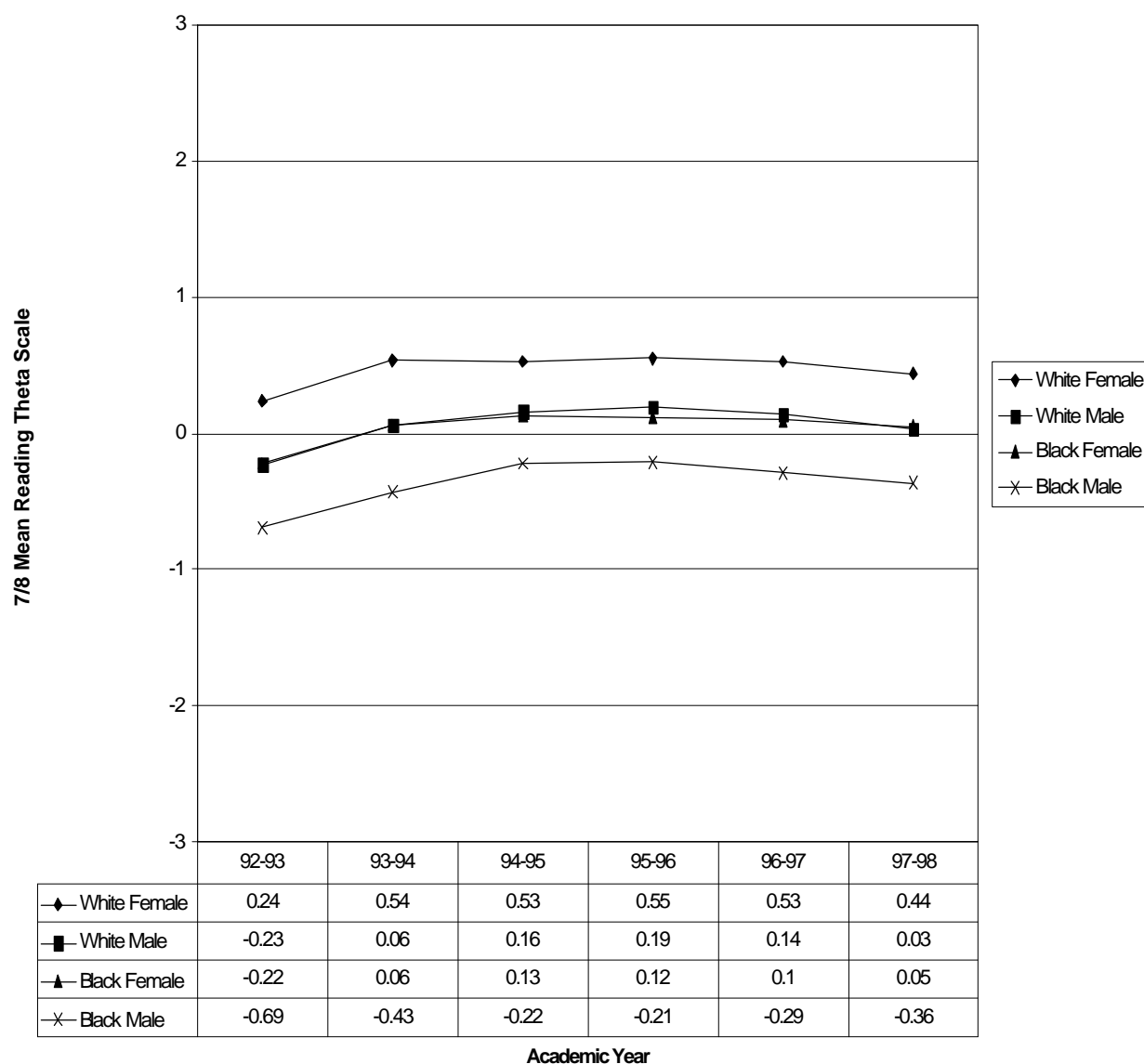
Figure 8. KIRIS Reading Theta for Grade 7/8 by Gender and Race, 1993-1998¹³¹³Data Table was added to Figure 8 by KDE.

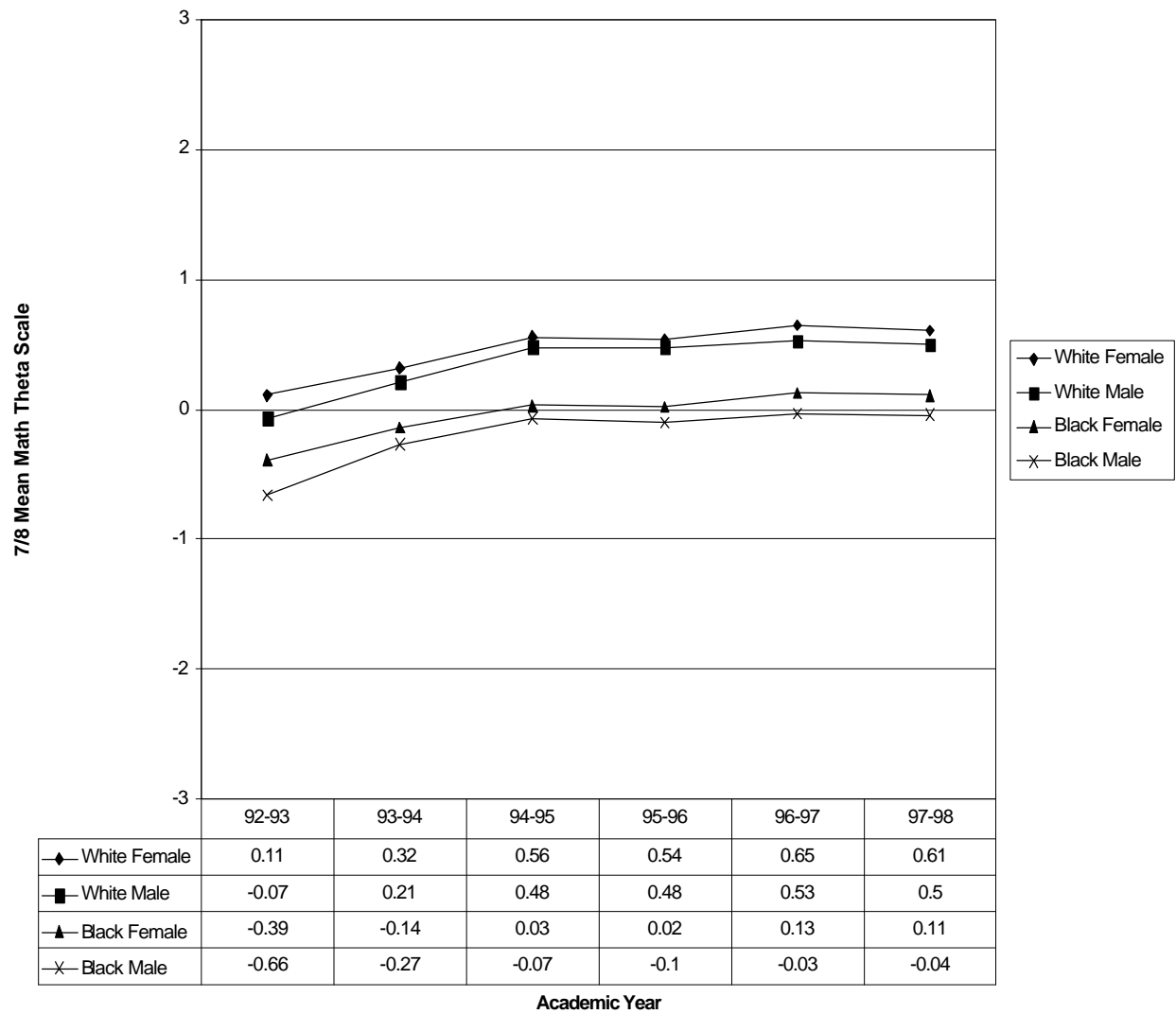
Figure 9. KIRIS Mathematics Theta for Grade 8 by Gender and Race, 1993-1998¹⁴¹⁴ Data Table was added to Figure 9 by KDE.

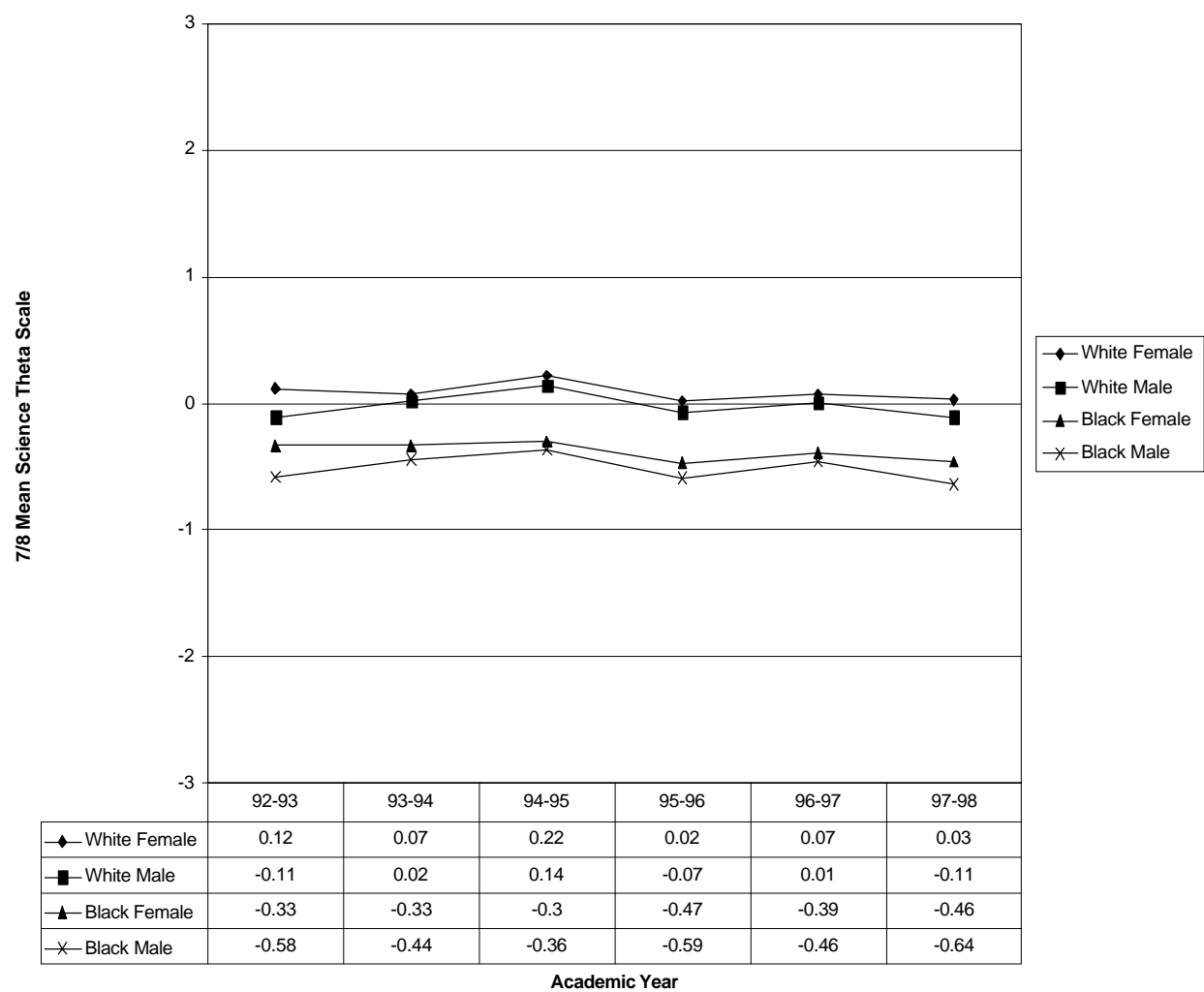
Figure 10. KIRIS Science Theta for Grade 7/8 by Gender and Race, 1993-1998¹⁵¹⁵ Data Table was added to Figure 10 by KDE.

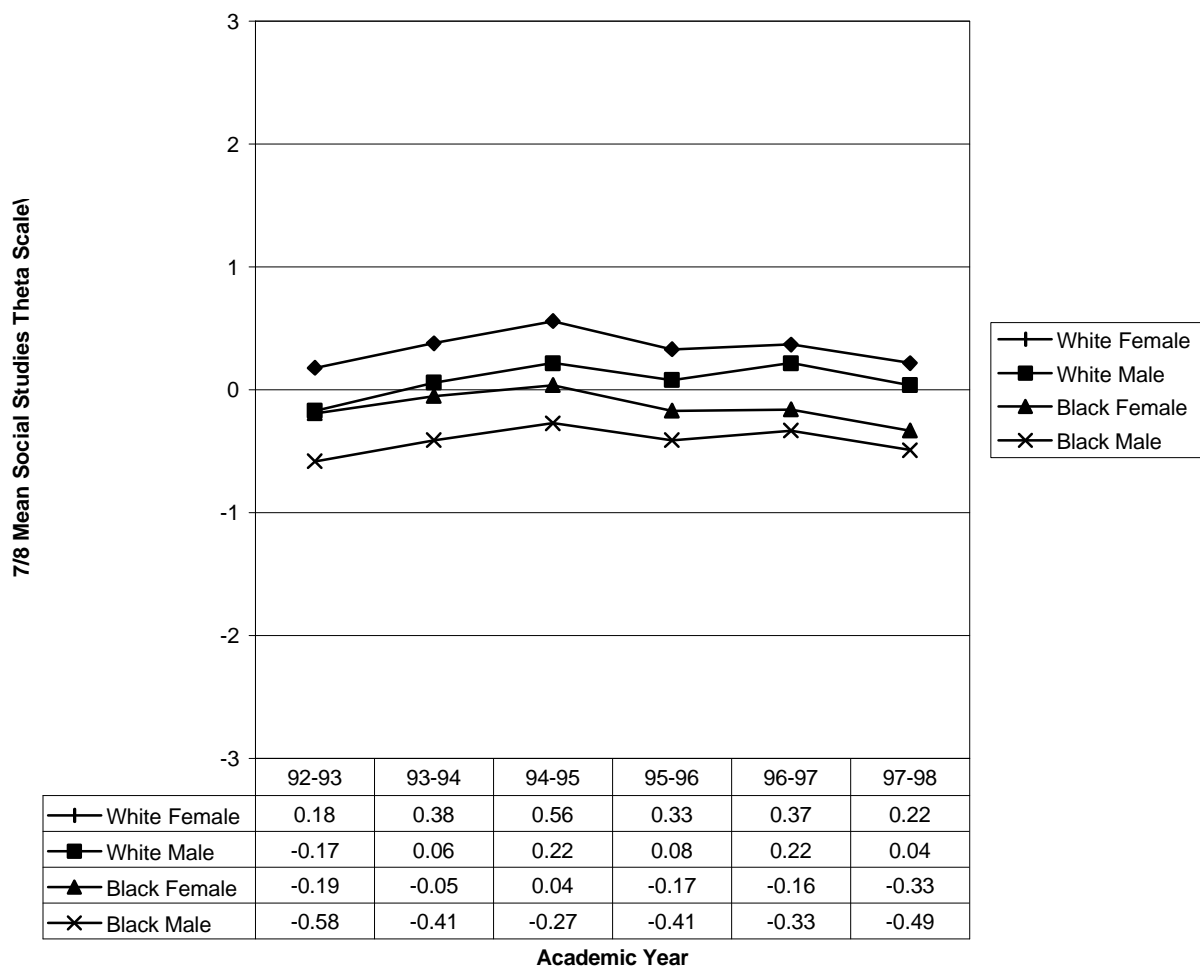
Figure 11. KIRIS Social Studies Theta for Grade 8 by Gender and Race, 1993-1998¹⁶¹⁶ Data Table was added to Figure 11 by KDE.

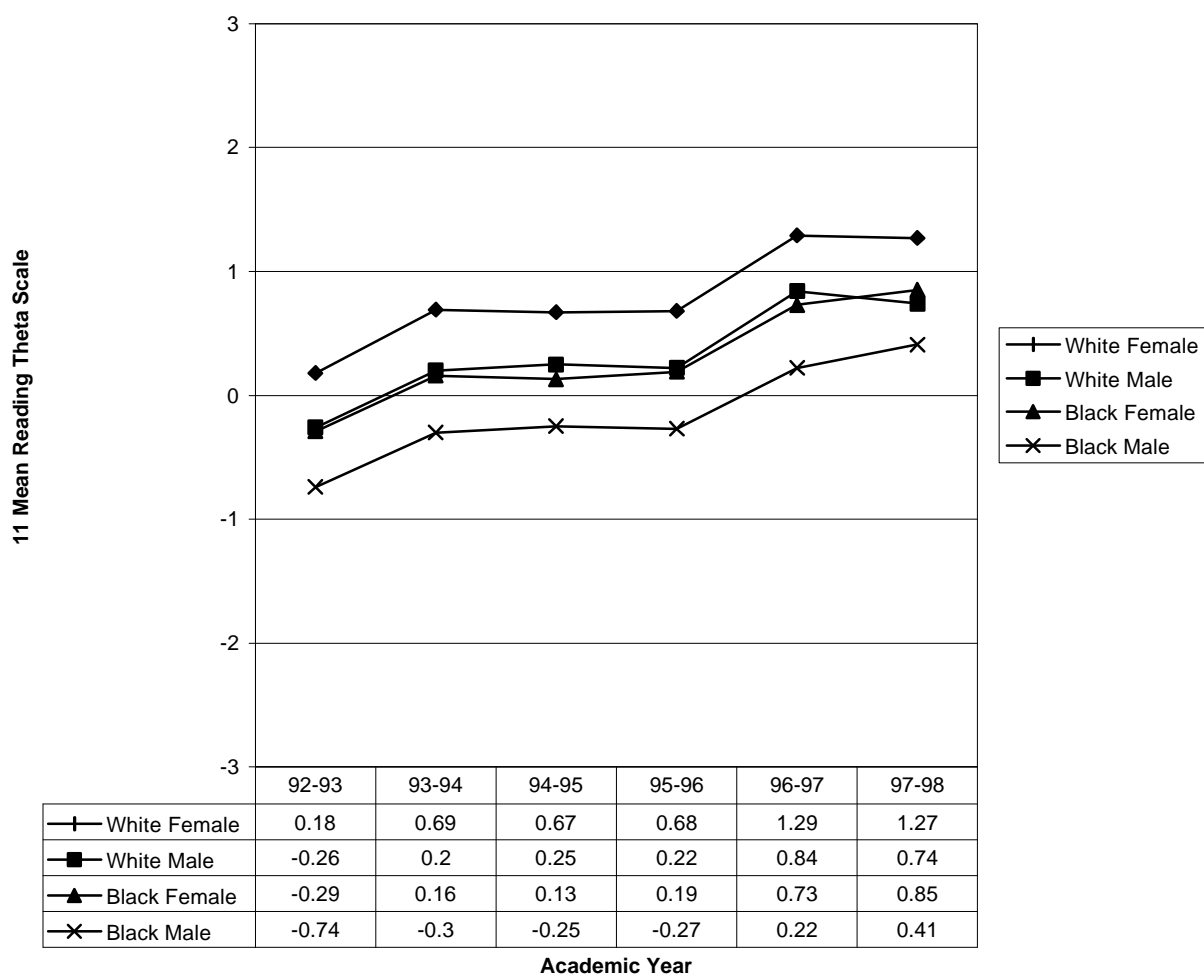
Figure 12. KIRIS Reading Theta for Grade 11/12 by Gender and Race, 1993-1998¹⁷¹⁷Data Table was added to Figure 11 by KDE.

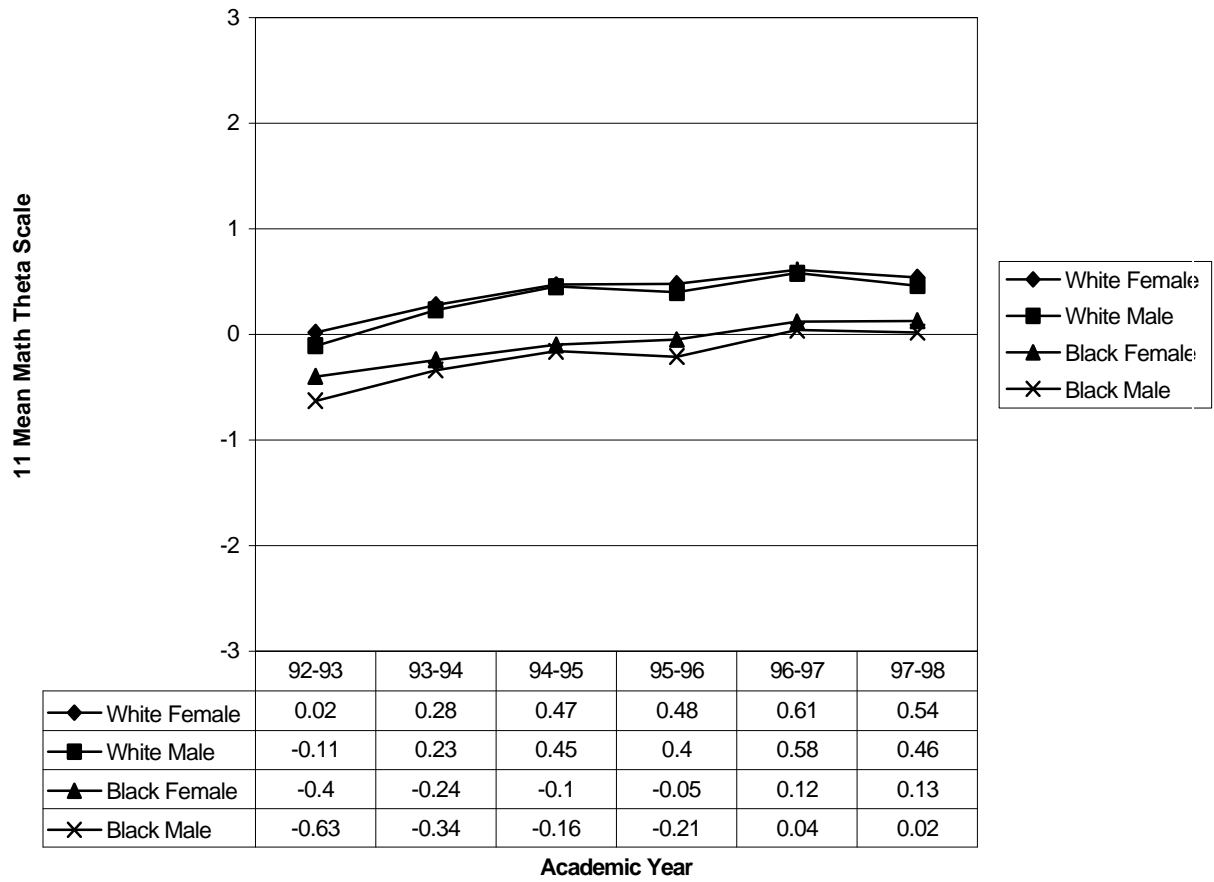
Figure 13. KIRIS Mathematics Theta for Grade 11/12 by Gender and Race, 1993-1999¹⁸¹⁷ Data Table was added to Figure 13 by KDE.

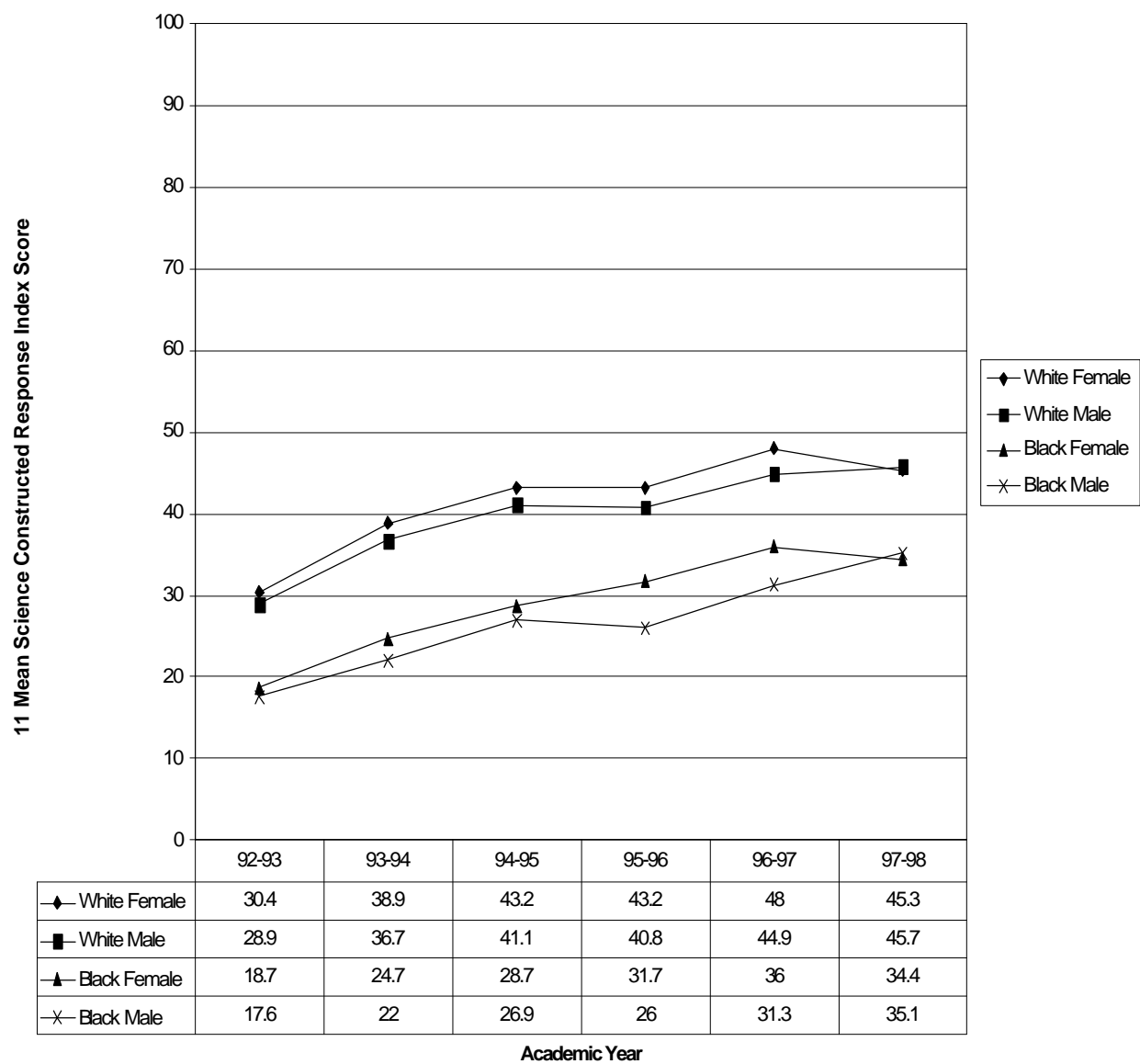
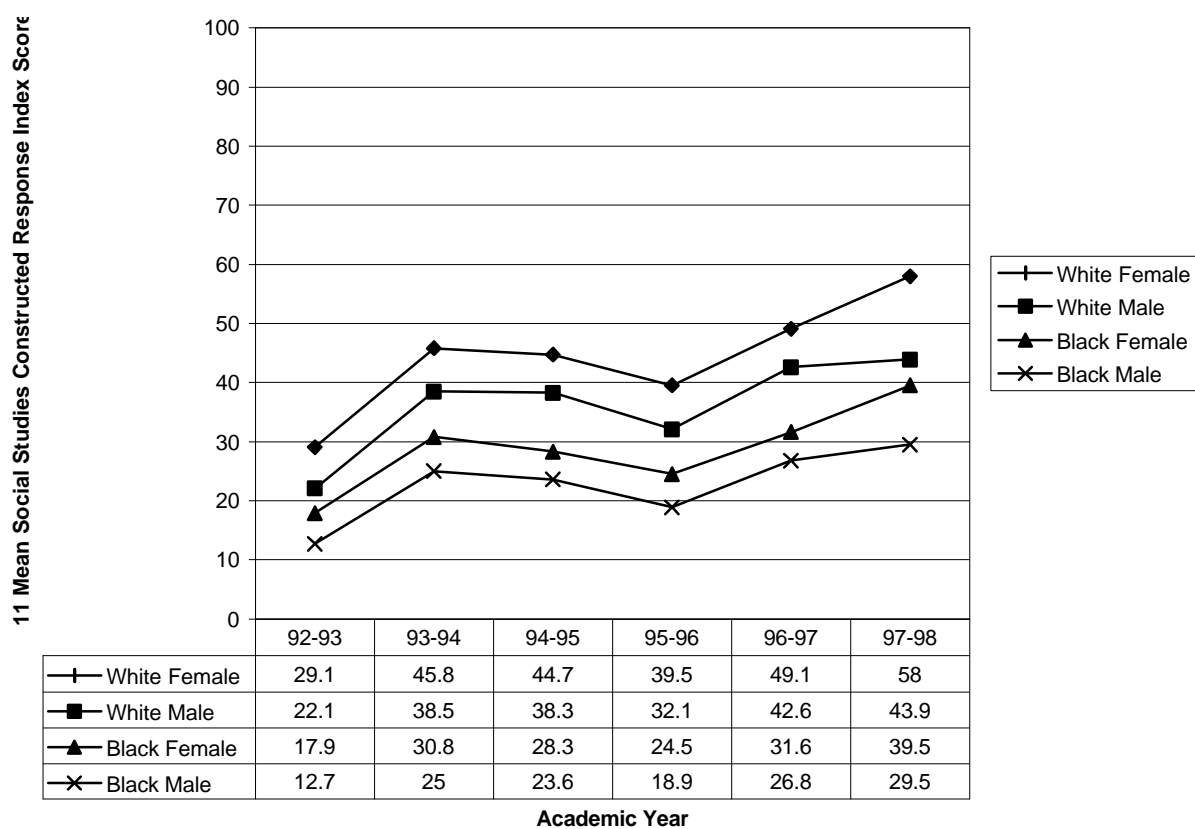
Figure 14. KIRIS Science Theta for Grade 11/12 by Gender and Race, 1993-1998¹⁹¹⁹ Data Table was added to Figure 14 by KDE.

Figure 15. KIRIS Social Studies Theta for Grade 11/12 by Gender and Race, 1993-1998²⁰²⁰ Data Table was added to Figure 15 by KDE.

attending public school in an accountability grade through KIRIS (i.e., the entire population), statistical significance is irrelevant. Any performance difference that appears in the data is a “true” difference in the population. The more important question to ask is how large are the observed differences.

Performance Differences: How Big are They?

Social scientists and physical scientists usually describe significance in statistical terms; however, social scientists must also assess the social significance of their findings. Because the KIRIS assessment includes the entire population, each significant difference in performance is an actual difference. However, significance in this sense does not indicate whether the observed differences are large enough to suggest implementation of policies to reduce them. Policy makers need to know how big the differences are between the groups.

While the information in the preceding figures provide a good pictorial view of performance differences over time, they can offer only a limited assessment of the size of performance differences between males and females and between blacks and whites. Social scientists have developed several statistical measures in the attempt to measure the size of experimental and quasi-experimental effects (Bjorgvinsson and Kerr 1995; Cohen 1977; Friedman 1968; Haase, Ellis and Ladany 1989; McGraw and Wong 1992; Sechrest and Yeaton 1982). One promising effect size measure from a policy for use in policymaking is the Common Language Effect Size Statistic (CLES).

The Common Language Effect Size Statistic converts the effect size into a probability. That is, it gives one the probability that a score taken at random from one distribution will be higher than a score taken from another distribution (McGraw and Wong 1992). Let us take a concrete example. In Table 3 are the Common Language Effect Size Statistics for male-female performance differences. In the first column, the Common Language Effect Size Statistics for 4th grade reading for the difference between males and females using the CRI scale is .42. This means, that if one drew a random 4th grade male student and a random 4th female student, the male would outscore the female in reading 42% of the time. When the theta scale is used, the statistic suggests that males would outscore females only 41% of the time. Of course, if there were really no difference between males and females, we would expect that males would outscore females 50% of the time and females would outscore males 50% of the time. Thus, the closer the Common Language Effect Size Statistic is to 50%, the closer the two groups are to having equal scores.

Data in Table 3 indicate that male and female student scores are nearly equal in mathematics and science. However, the differences between males and females are larger in reading and social studies. In both of those subjects males are likely to be outscored by females at least 5% more frequently than would be expected if males and females were equal. As we move up in grade level, the differences between males and females on the reading tests increase. In grade 11 males are likely to outscore females only a third of the time. It is unclear whether this

difference is due to age or cohort effects. Still, the effect size statistics shown in Table 3 suggest that gender accounts for little of the test score differences in math and science and less than 10% of the variance in KIRIS reading and social studies tests (McGraw and Wong 1992).

Table 3. Common Language Effect Size Statistics for Male-Female Performance Differences

	Grade 4		Grade 5		Grade 7		Grade 8		Grade 11	
	CRI	Theta	CRI	Theta	CRI	Theta	CRI	Theta	CRI	Theta
Reading	.42	.41			.39	.36			.36	.34
Mathematics			.47	.47			.47	.46	.48	.48
Science	.49	.49			.45	.45			.50	.50
Social Studies			.43	.43			.44	.44	.41	.39

In Table 4 are the Common Language Effect Size Statistics for the differences between whites and African-Americans. While were less likely to outscore females, a randomly sampled white student is at least 10 percent more likely to outscore a randomly sampled black student in every subject area at every grade level. As noted earlier, the difference is particularly large in mathematics and science, where whites are likely to outscore blacks slightly more than two thirds of the time. Unlike the gender difference, the racial performance difference is smallest at Grade 11. This smaller difference could be due to a disproportionate number of black students dropping out, or it could signal the fact that performance differences are increasing over time (Calabrese and Poe 1991).²¹ Again, McGraw and Wong (1992) would suggest that race explains approximately 10 percent of the variance in all four KIRIS subtests.

Performance Differences: Increasing or Decreasing?

In Table 5, are the performance differences between males and females measured in 1992-93 and 1997-98. These performance differences are nearly identical. In general, the probability of males being outscored by females has slightly decreased in most areas for most grades; however, these decreases are only 1 or 2 percent. The only exception is in 8th grade social studies where there is a decrease in the performance difference of nearly 5 percent. An examination of data in Figure 11 suggests that one reason for this gender decrease may be an increase in race differences. The observed difference between white male and female scores decreases through increases made by white males while the difference between black male and female scores decreases through a decrease in black female test scores.

¹⁶While dropout rates are kept by race at the school and school district level, at this time, the Kentucky Department of Education does not collect that information from the local agencies. However, movement appears to be underway to make such information available in the near future.

Table 4. Common Language Effect Size Statistics for White/African-American Performance Differences

	Grade 4		Grade 5		Grade 7		Grade 8		Grade 11	
	CRI	Theta	CRI	Theta	CRI	Theta	CRI	Theta	CRI	Theta
Reading	.59	.62			.61	.64			.60	.61
Mathematics			.67	.70			.67	.69	.64	.65
Science	.65	.67			.67	.70			.64	.68
Social Studies			.62	.64			.64	.67	.62	.63

Table 5. Change in the Common Language Effect Size Statistics for Male/Female Performance Differences, 1992-93 to 1997-98

		Grade 4 and 5			Grade 7 and 8			Grade 11/12		
		1992-93	1997-98	Change	1992-93	1997-98	Change	1992-93	1997-98	Change
Reading	CRI	.42	.42	0	.39	.39	0	.41	.36	Increase
	Theta	.41	.41	0	.37	.36	Increase	.37	.34	Increase
Mathematics	CRI	.47	.47	0	.48	.47	Increase	.50	.48	Increase
	Theta	.46	.47	Decrease	.44	.46	Decrease	.46	.48	Decrease
Science	CRI	.48	.49	Decrease	.45	.45	0	.48	.50	Decrease
	Theta	.47	.49	Decrease	.43	.45	Decrease	.47	.50	Decrease
Social Studies	CRI	.46	.43	Increase	.42	.44	Decrease	.43	.41	Increase
	Theta	.44	.43	Increase	.40	.44	Decrease	.41	.39	Increase

Increases in the difference between male and female performance scores were not only rare but also small. Most often these differences increased by only 1 or 2 percent. The exception is the reading difference between males and females at Grade 11. The probability that a randomly sampled female will outscore a randomly sampled male has increased by 4 to 5 percent. An examination of the data in Figure 15, suggests that the increase is due to white males, whose scores did not increase at the same rate as other groups over the last half of Cycle 3.

Changes in the size of the performance difference between African-American and white students are presented in Table 6. Unlike the gender performance difference, this analysis shows that most performance differences between white and African-American students widened slightly over the last six years. In 8th grade science and social studies the probability that whites would outscore African American students increased by 5 percent. In 1992-93 white eighth graders in social studies were likely to outscore their black counterparts 61 percent of the time. By 1997-98 their probability of outscoring black students had risen to 67 percent.

Table 6. Change in the Common Language Effect Size Statistics for White/African-American Performance Differences, 1992-93 to 1997-98.

		Grade 4 and 5			Grade 7 and 8			Grade 11/12		
		1992-93	1997-98	Change	1992-93	1997-98	Change	1992-93	1997-98	Change
Reading	CRI	.61	.59	Decrease	.60	.61	Increase	.61	.60	Decrease
	Theta	.62	.62	0	.63	.64	Increase	.62	.61	Decrease
Mathematics	CRI	.64	.67	Increase	.64	.67	Increase	.63	.64	Increase
	Theta	.66	.70	Increase	.66	.69	Increase	.64	.65	Increase
Science	CRI	.63	.65	Increase	.62	.67	Increase	.64	.64	0
	Theta	.65	.67	Increase	.63	.70	Increase	.66	.68	Increase
Social Studies	CRI	.63	.62	Decrease	.59	.64	Increase	.61	.62	Increase
	Theta	.60	.64	Increase	.61	.67	Increase	.62	.63	Increase

To summarize, performance differences are occurring between males and females and between blacks and whites. Gender and race effects taken together, explain less than 10 percent of the variance in each of the subject area test scores. Gender performance differences, for the most part, appear to be slowly decreasing except for 11th grade reading. Race differences, on the other hand, appear to be increasing slowly in math at all grades, in science and social studies in the 4th, 5th and 11th grades, and in 7th grade reading. Moreover, 7th grade science and 8th grade math have seen a more substantial increase in the white/African-American performance difference over the same period.

EXPLORATIONS INTO PERFORMANCE DIFFERENCES BY GENDER AND RACE

Although there are performance differences evident within the KIRIS data between males and females and between whites and African-Americans, the size of the performance differences and the small amount of variance they explain do not necessarily imply that the KIRIS assessment or the accountability system is biased. The Bias Review Committee reviews the KIRIS assessment extensively during the test construction phase each year. Items judged by the Bias Review Committee to cause disadvantage to student subgroups are dropped from the test or modified to remove the potential source of bias. In addition, examination of the available data from the National Assessment of Educational Progress (NAEP) and the ACT (Figures 16 through 23) reveals similar performance differences between males and females and between white and African-American students (Education Week reports that the race difference decreased from 1970 to 1988 and has been increasing ever since [Olson 1997]).²² These differences are roughly comparable to those found on KIRIS.

Given the small amount of variance explained by gender and race as measured by the Common Language Effect Size statistics, the existing safeguards, the continuing research to detect bias, and the comparability of results from other testing programs, bias in the KIRIS tests may account for a minor proportion of the differential performances (Boone 1997, Collins-Eaglin and Karabenick, 1993, Fernanda, Morais and Neves 1995, Frisby 1998, Jencks and Phillips 1998, Samejiima 1969). In addition to continuing research on performance biases, a major focus needs to be on gathering the individual, community, school, or district level data to investigate other factors that might contribute to the observed differences. Much of the data necessary for investigating these possible effects thoroughly has not always been available to the Kentucky Department of Education. The following three exploratory studies reflect ongoing research into observed performance differences. These exploratory studies are: 1) an examination of the effects of gender and racial concentration on subject area performance differences, 2) a qualitative examination of schools with low performance differences between whites and African-Americans, and 3) a quantitative examination of the relationship between a school's accountability index score, its gender and racial mix and socioeconomic status.

²² NAEP and ACT data also show that the racial performance difference observed in Kentucky is smaller than the national difference in performance between whites and African-Americans. However, the smaller difference occurs because Kentucky whites score lower than the national white average while Kentucky African-Americans score slightly higher than African-American students nationwide. It should be noted that the NAEP assessment is only given in a sample of Kentucky's schools on a voluntary basis, which may affect the size of the performance differences shown. Furthermore, only students planning to go to college tend to take the ACT test; thus, the differences shown are not necessarily representative of the Commonwealth's entire student body. Still, these data currently provide the best comparison measures for the KIRIS assessment data and do show similar patterns.

Figure 16. United States and Kentucky 4th Grade NAEP Reading Scores by Gender

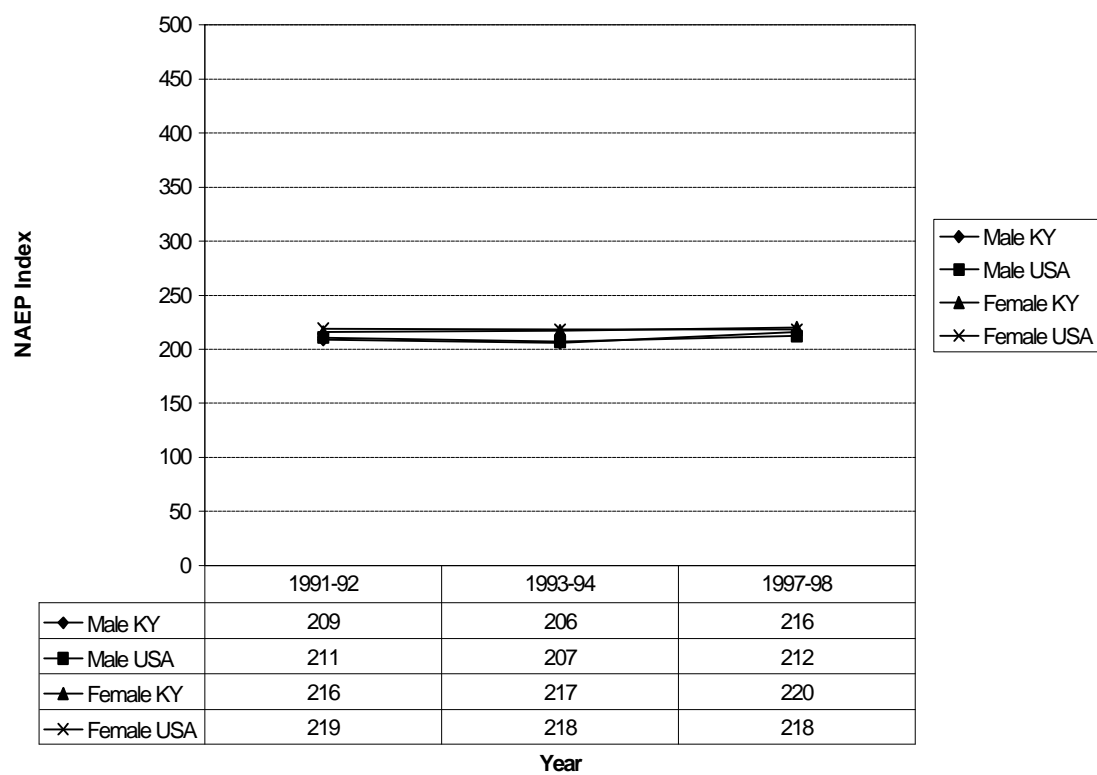


Figure 17. United States and Kentucky 4th Grade NAEP Mathematics Scores by Gender

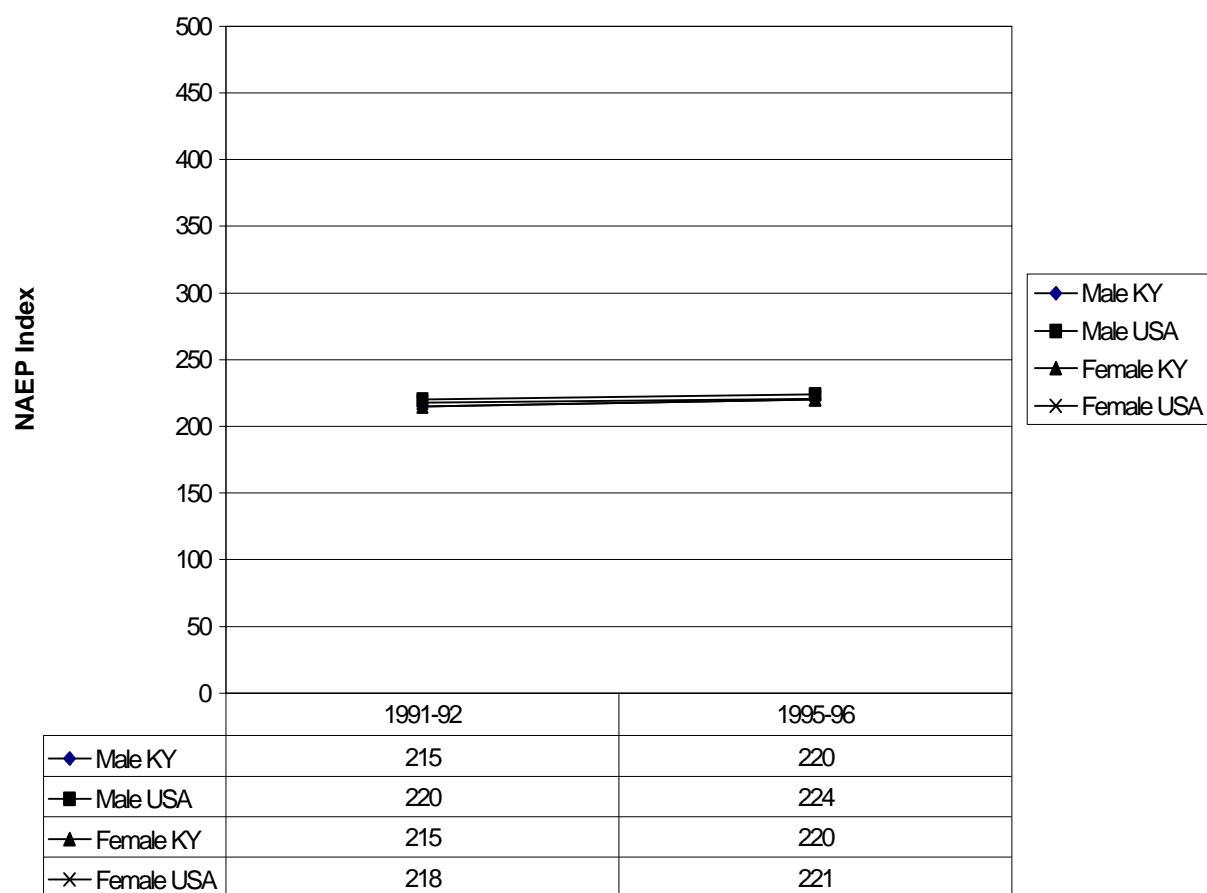


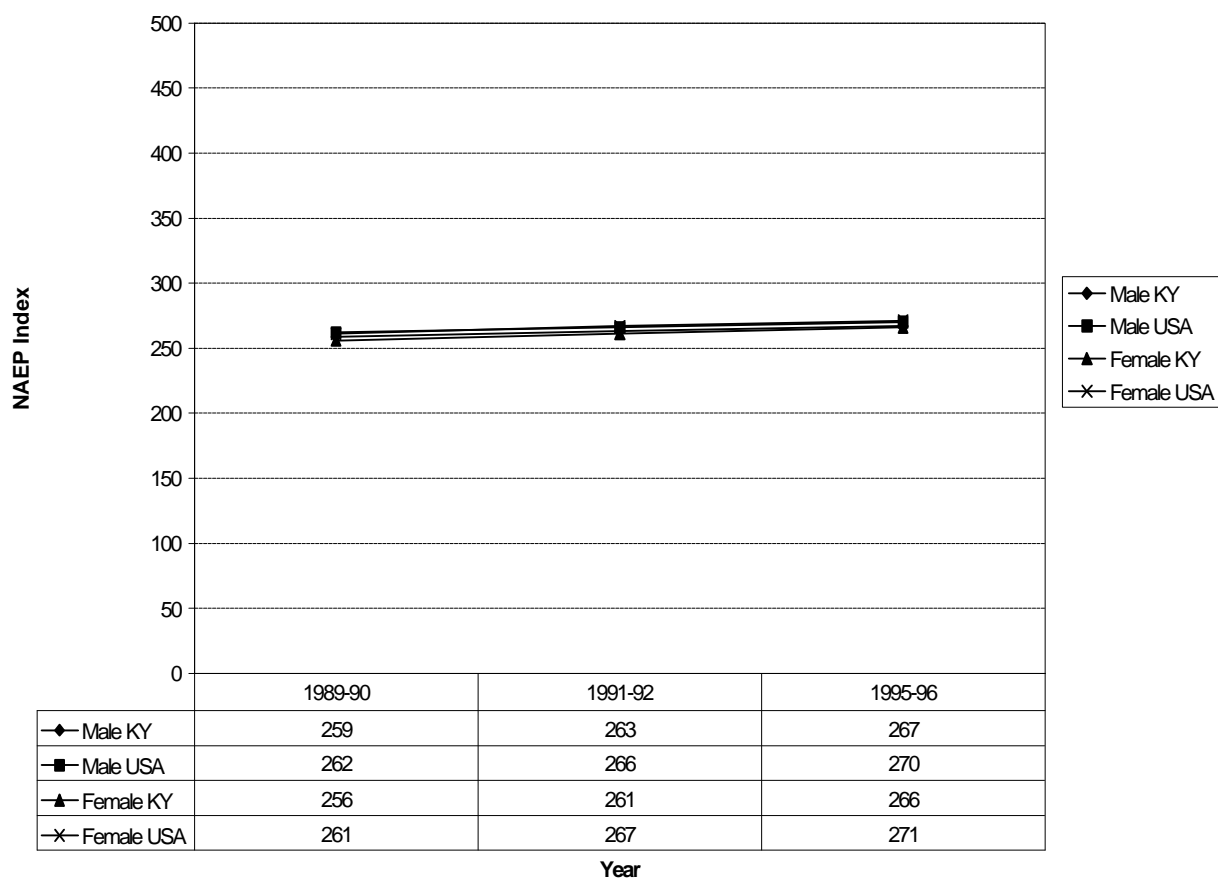
Figure 18. United States and Kentucky 8th Grade NAEP Mathematics Scores by Gender

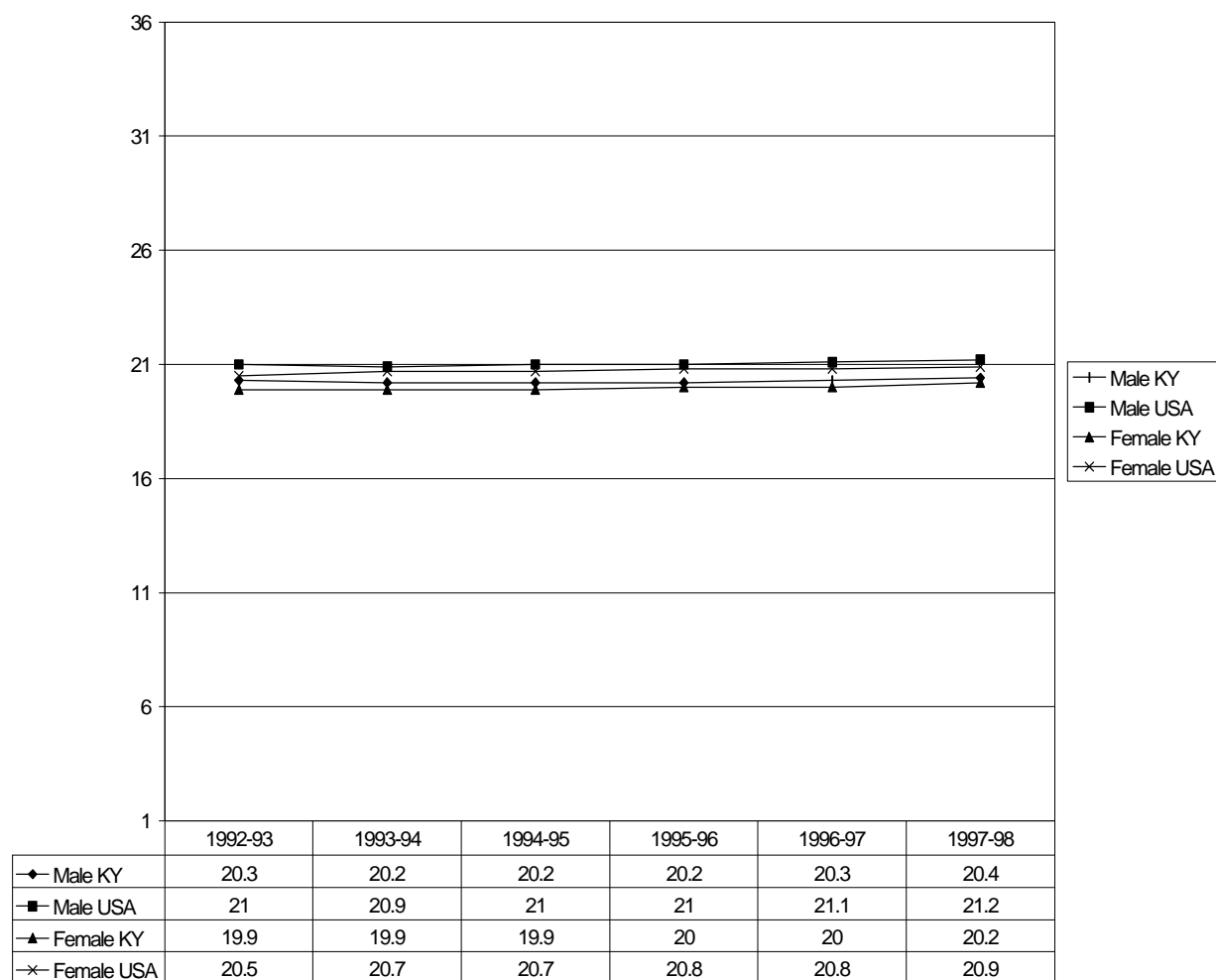
Figure 19. Comparison of United States and Kentucky ACT Results by Gender, 1993-1998²³²³ Data Table was added to Figure 19 by KDE.

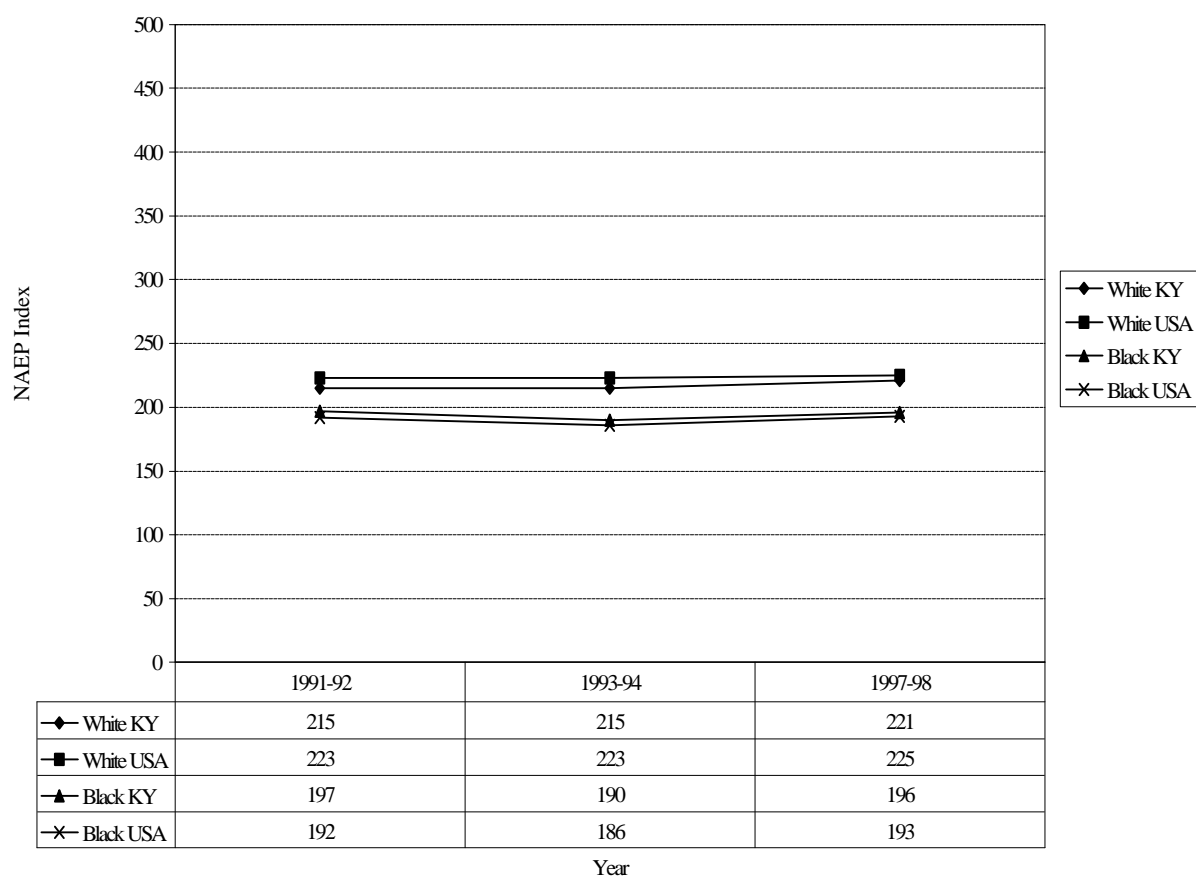
Figure 20. United States and Kentucky 4th Grade NAEP Reading Scores by Race²⁴²⁴ Data Table was added to Figure 20 by KDE.

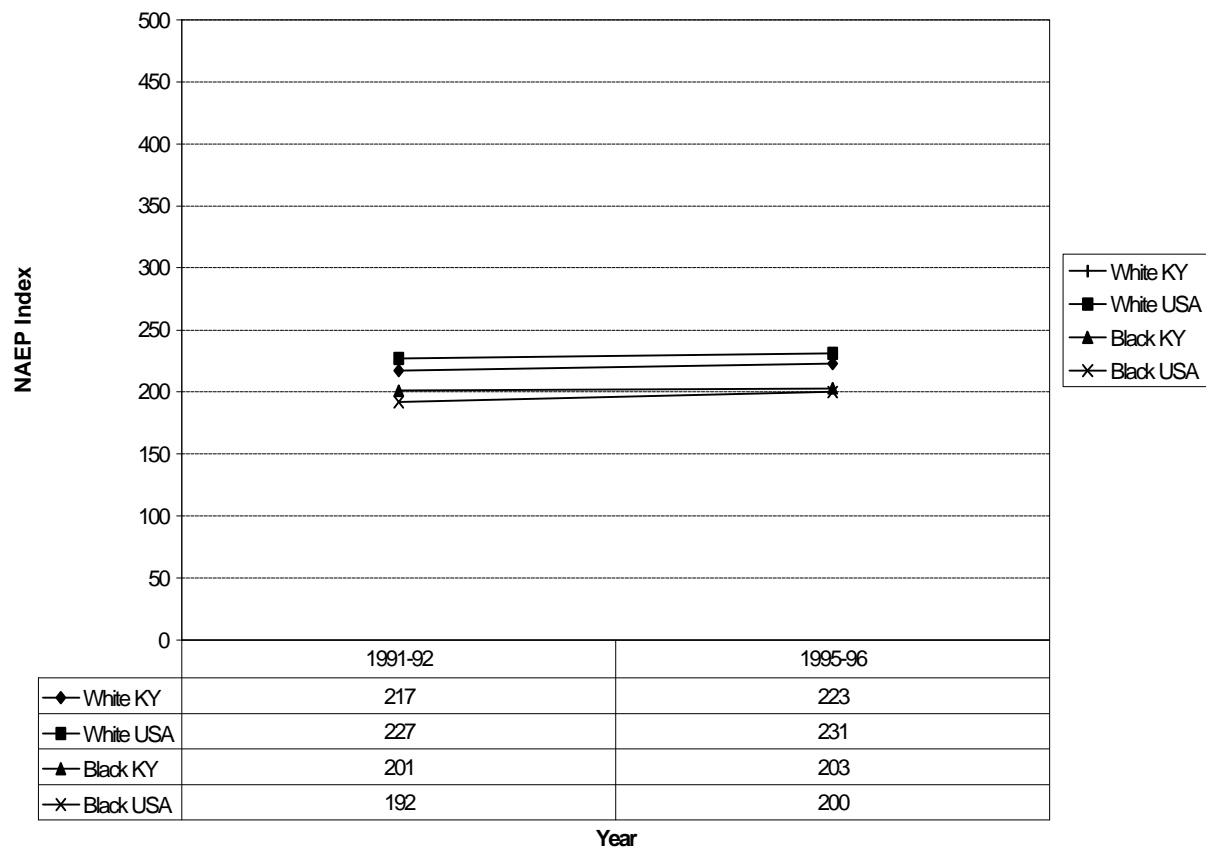
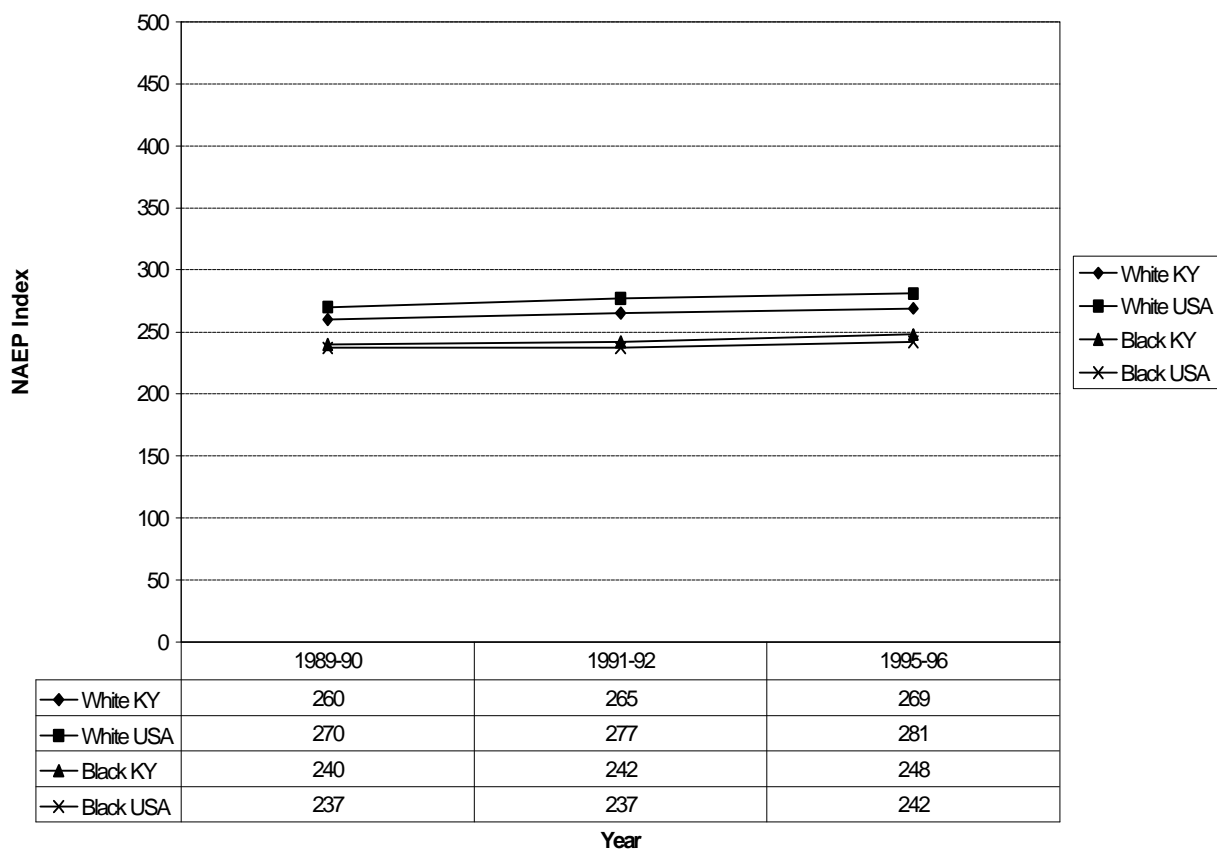
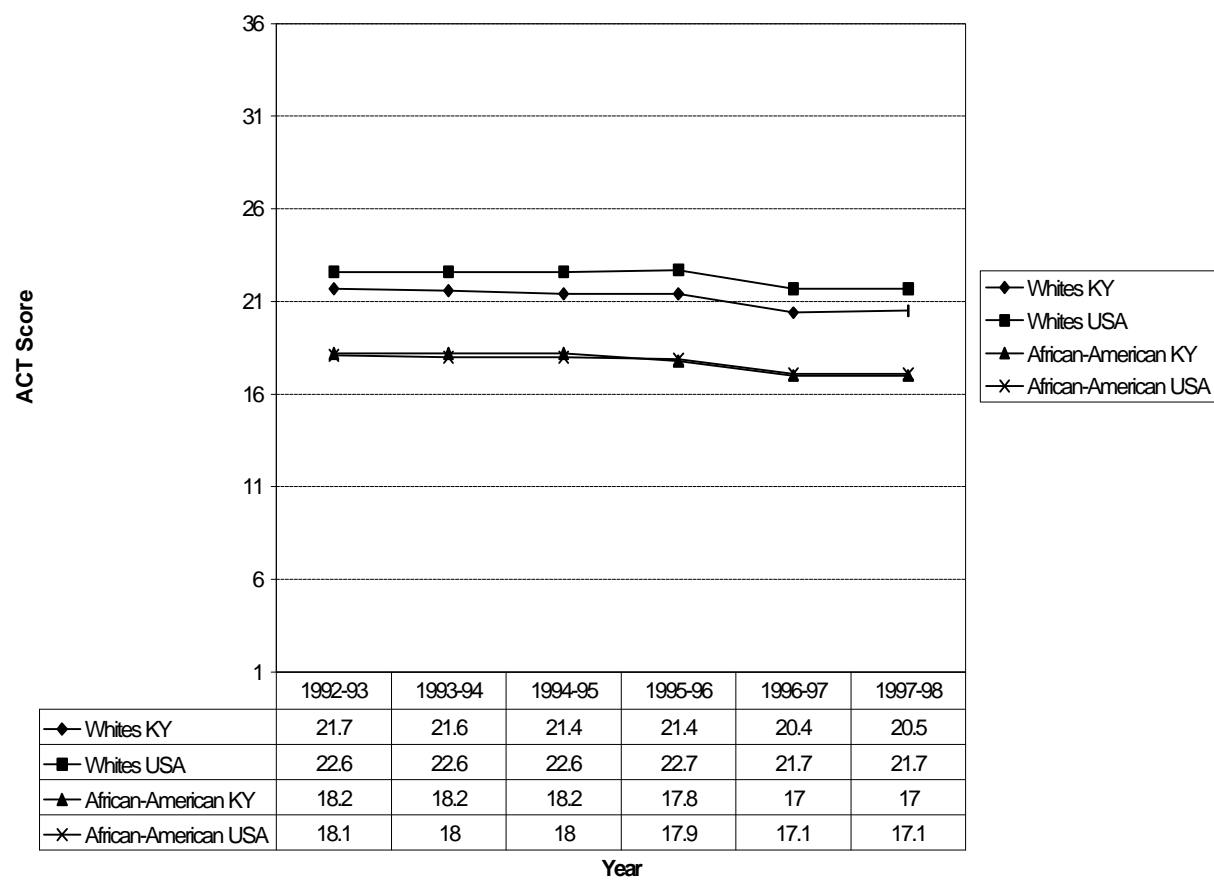
Figure 21. United States and Kentucky 4th Grade NAEP Mathematics Scores by Race²⁵²⁵ Data Table was added to Figure 21 by KDE.

Figure 22. United States and Kentucky 8th Grade NAEP Mathematics Scores by Race²⁶

²⁶ Data Table was added to Figure 22 by KDE.

Figure 23. Comparison of United States and Kentucky ACT Results by Race, 1993-1998²⁷

²⁷ Data Table was added to Figure 23 by KDE.

Exploratory Study 1: The Geographic Concentration of Gender and Race in Kentucky

The term geographic concentration can be simply defined as the idea that certain subgroups of the population (i.e., minorities) are concentrated differentially within and across space. This concept is an important one in social science because research shows that the concentration of subgroups of population within and across space can affect their experiences and their life chances (Hays 1989; Wilkinson 1991; Wilson 1987). For example, Wilson (1987) argues that as lower class blacks become geographically concentrated in the inner city, their life chances decline.

The term minority is usually defined in terms of a numerical minority, i.e., an identifiable subgroup whose numbers comprise less than 50 percent of the population. Since the 1960s this term has been redefined to include categories of individuals marginalized not by numbers but by decision-making power. Gender is a good example of this type of minority. As seen in Table 1 there are more females than males in the KIRIS accountability grades. However, traditional sex roles in the Commonwealth, as in the United States as a whole, have historically given men and women very different educational and career opportunities (. AAUW 1992, Catsambis 1994, Kanter 1977, Kleinfeld 1998).

Not all minorities are geographically concentrated. In Kentucky schools, females are not physically or spatially segregated from males. Thus, it is doubtful that any comparison of male and female test score differences will vary by geographic concentration. However, unlike gender, racial minorities are not evenly geographically spread throughout the Commonwealth's 120 counties but instead are found primarily in the major cities and near the military bases in the state. Moreover, new research by Orfield and Yun (1999) suggests that schools in the American South may be resegregating, increasing the concentration of blacks and other minority students in certain geographic areas. The uneven concentration of minority students may profoundly influence the size of performance differences between African-Americans and whites.

Several levels of social organization can be examined for geographic concentration. For example in education, one can measure geographic concentration at the state level, county level, school district level, school level, and classroom level. Regarding the KIRIS assessment, two of these possible levels of measurement, school district and school, are included by KDE with students' assessment scores. Occasionally, state and federal agencies make data on other levels of measurement available which categorize districts into rural and urban. The Census Bureau defines urban as comprising all places within urbanized areas (one or more places and the adjacent densely surrounding territory) and places with 2,500 or more persons outside urbanized areas (NCES 1994). Rural, then, is any area not urban.

The Geographic Concentration of Gender

To examine potential gender performance difference in test scores based on geographic concentration, the percentage of females enrolled in each school district and each school was

calculated. Next the school district median female enrollment and each school median female enrollment were calculated. In both cases the median district and the median school had a student body that averaged 49% female. Three tables were then developed using this information. The first, Table 7, presents the Common Language Effect Size statistics for two groups of students: those who attend school in districts with higher than the median percentage of female students enrolled and those who attend school in districts with equal to or lower percentages of female enrollment. Table 8 presents the same information, except that this time the students are grouped by the median percentage of females enrolled in their schools. The third table, Table 9, presents the Common Language Effect Size Statistics for students based on both the median percentage of women in each district and the median percentage of women in each school.

Data in Tables 7, 8, and 9, indicate that while there is a disparity between male and female student performance, any form of geographic concentration causes little of this disparity. The variation in the performance difference between males and females by school district and/or by school is about 1 to 2 percentage points and appears more often in Grades 7, 8, and 11. Data in Table 9, indicate that the variation between male and female student differences is larger among the four groups of students at the 11th grade; however, the effect size statistic never varies more than 4 percent due to the gender concentration in school districts or schools.

Geographic Distribution of Race

To examine the possibility that students in school districts with higher percentages of minority students enrolled differ in their achievement from students in school districts with lower minority percentages, the students in each grade tested were divided into two groups based on whether their district was above or below the median percentage minority enrollment (3%).²⁸ Common Language Effect Size Statistics were then calculated for the performance differences for students in each division. The results of these analyses are presented in Tables 10, 11, and 12.

Unlike gender, the size of the performance difference between white and African-American students varies with the racial mix of the district and school. First, the probability that whites outscore African-American students is larger for students in the districts with above the 3% minority enrollment than it is for students in districts with below the state average (See Table 10). In Grades 4 and 5, on average, whites are likely to outscore blacks by nearly 10% more of the time if the students attend a district with above a 3% minority population. A disparity of the same magnitude exists for Grades 7 and 8. Whites are nearly 10 percent more likely to have higher scores than blacks in school districts with above 3 percent minority populations than are whites in below state average minority districts. This disparity is not nearly so evident at Grade

²⁸The school and district level analysis uses percent of minority enrolled in the accountability grades instead of percent of African-Americans enrolled; however, these two measures are comparable given the small number of other minority students.

Table 7. Difference Between Male and Female Means Measured in Common Language Effect Size Statistics for Students in School Districts with Above and Below the Median State Percent of Female Enrollment

		Districts with > Kentucky Median Female Enrollment	Districts with < = Kentucky Median Female Enrollment
Grade 4 and 5	Reading CRI	.42	.42
	Reading Theta	.41	.41
	Math CRI	.47	.47
	Math Theta	.47	.47
	Science CRI	.49	.49
	Science Theta	.49	.49
	Social Studies CRI	.43	.43
	Social Studies Theta	.43	.43
Grade 7 and 8	Reading CRI	.39	.39
	Reading Theta	.36	.35
	Math CRI	.47	.47
	Math Theta	.46	.46
	Science CRI	.45	.45
	Science Theta	.45	.44
	Social Studies CRI	.45	.44
	Social Studies Theta	.44	.43
Grade 11	Reading CRI	.37	.35
	Reading Theta	.35	.33
	Math CRI	.49	.48
	Math Theta	.48	.47
	Science CRI	.51	.50
	Science Theta	.51	.50
	Social Studies CRI	.41	.40
	Social Studies Theta	.40	.38

Table 8. Difference Between Male and Female Means Measured in Common Language Effect Size Statistics for Students in Schools with Above and Below the Median State Percent of Female Enrollment

		Schools with > Kentucky Median Female Enrollment	Schools with < = Kentucky Median Female Enrollment
Grade 4 and 5	Reading CRI	.42	.42
	Reading Theta	.41	.41
	Math CRI	.47	.47
	Math Theta	.47	.47
	Science CRI	.50	.49
	Science Theta	.49	.49
	Social Studies CRI	.43	.43
	Social Studies Theta	.43	.43
Grade 7 and 8	Reading CRI	.40	.39
	Reading Theta	.36	.35
	Math CRI	.47	.47
	Math Theta	.46	.46
	Science CRI	.46	.45
	Science Theta	.45	.44
	Social Studies CRI	.45	.45
	Social Studies Theta	.44	.44
Grade 11	Reading CRI	.36	.35
	Reading Theta	.35	.34
	Math CRI	.49	.48
	Math Theta	.47	.47
	Science CRI	.51	.50
	Science Theta	.51	.50
	Social Studies CRI	.41	.40
	Social Studies Theta	.40	.39

Table 9 Difference Between Male and Female Means Measured in Common Language Effect Size Statistics for Students in School Districts and Schools with Above and Below the Median State Percent of Female Enrollment

		Districts with > Kentucky Median Female Enrollment		Districts with < = Kentucky Median Female Enrollment	
		Schools with > Kentucky Median Female Enrollment	Schools with < = Kentucky Median Female Enrollment	Schools with > Kentucky Median Female Enrollment	Schools with < = Kentucky Median Female Enrollment
Grade 4 and 5	Reading CRI	.41	.43	.42	.42
	Reading Theta	.40	.42	.41	.41
	Math CRI	.48	.47	.46	.48
	Math Theta	.47	.47	.46	.47
	Science CRI	.49	.49	.50	.49
	Science Theta	.49	.49	.50	.49
	Social Studies CRI	.44	.43	.42	.44
	Social Studies Theta	.43	.43	.42	.43
Grade 7 and 8	Reading CRI	.40	.39	.40	.39
	Reading Theta	.36	.35	.36	.35
	Math CRI	.47	.47	.47	.47
	Math Theta	.46	.46	.46	.46
	Science CRI	.46	.45	.45	.45
	Science Theta	.45	.44	.44	.44
	Social Studies CRI	.45	.45	.44	.44
	Social Studies Theta	.44	.45	.44	.43
Grade 11	Reading CRI	.37	.37	.35	.35
	Reading Theta	.35	.36	.34	.32
	Math CRI	.49	.50	.48	.47
	Math Theta	.48	.49	.47	.46
	Science CRI	.51	.52	.51	.50
	Science Theta	.51	.52	.51	.49
	Social Studies CRI	.41	.42	.40	.40
	Social Studies Theta	.40	.41	.39	.37

Table 10. Common Language Effect Size Statistics for the Difference between White and African American Students in School Districts with Above and Below the State Percent Minority Enrollment Average

		Districts with > Kentucky Median Minority Enrollment	Districts with < = Kentucky Median Minority Enrollment
Grade 4 and 5	Reading CRI	.61	.51
	Reading Theta	.64	.53
	Math CRI	.70	.59
	Math Theta	.72	.64
	Science CRI	.66	.55
	Science Theta	.69	.59
	Social Studies CRI	.64	.57
	Social Studies Theta	.67	.58
Grade 7 and 8	Reading CRI	.63	.58
	Reading Theta	.66	.59
	Math CRI	.70	.59
	Math Theta	.71	.60
	Science CRI	.69	.61
	Science Theta	.72	.60
	Social Studies CRI	.67	.57
	Social Studies Theta	.69	.57
Grade 11	Reading CRI	.62	.65
	Reading Theta	.63	.64
	Math CRI	.67	.64
	Math Theta	.68	.63
	Science CRI	.65	.61
	Science Theta	.70	.65
	Social Studies CRI	.64	.64
	Social Studies Theta	.65	.63

Table 11. Common Language Effect Size Statistics for the Difference between White and African American Students in Schools with Above and Below the State Percent Minority Enrollment Average

		Schools with > Kentucky Median Minority Enrollment	Schools with < = Kentucky Median Minority Enrollment
Grade 4 and 5	Reading CRI	.61	.54
	Reading Theta	.63	.55
	Math CRI	.69	.60
	Math Theta	.72	.64
	Science CRI	.65	.53
	Science Theta	.68	.57
	Social Studies CRI	.64	.60
	Social Studies Theta	.66	.62
Grade 7 and 8	Reading CRI	.63	.53
	Reading Theta	.66	.55
	Math CRI	.70	.60
	Math Theta	.71	.61
	Science CRI	.70	.57
	Science Theta	.72	.58
	Social Studies CRI	.67	.60
	Social Studies Theta	.69	.61
Grade 11	Reading CRI	.62	.60
	Reading Theta	.63	.60
	Math CRI	.67	.62
	Math Theta	.68	.60
	Science CRI	.65	.55
	Science Theta	.70	.61
	Social Studies CRI	.64	.64
	Social Studies Theta	.65	.61

Table 12. Common Language Effect Size Statistics for the Difference between White and African-American Students in School Districts and Schools with Above and Below the State Percent Minority Enrollment Average

		Districts with > Kentucky Median Minority Enrollment		Districts with < = Kentucky Median Minority Enrollment	
		Schools with > Kentucky Median Minority Enrollment	Schools with < = Kentucky Median Minority Enrollment	Schools with > Kentucky Median Minority Enrollment	Schools with < = Kentucky Median Minority Enrollment
Grade 4 and 5	Reading CRI	.61	.55	.46	.55
	Reading Theta	.64	.57	.49	.56
	Math CRI	.70	.63	.61	.59
	Math Theta	.72	.66	.66	.65
	Science CRI	.66	.59	.62	.50
	Science Theta	.68	.61	.62	.56
	Social Studies CRI	.64	.64	.58	.58
	Social Studies Theta	.66	.66	.56	.61
Grade 7 and 8	Reading CRI	.63	.44	.64	.58
	Reading Theta	.66	.47	.80	.58
	Math CRI	.70	.62	.56	.60
	Math Theta	.71	.64	.60	.60
	Science CRI	.70	.48	.87	.60
	Science Theta	.72	.54	.80	.59
	Social Studies CRI	.67	.69	.49	.58
	Social Studies Theta	.70	.69	.46	.59
Grade 11	Reading CRI	.62	.50	.72	.63
	Reading Theta	.63	.56	.72	.62
	Math CRI	.67	.73	.74	.61
	Math Theta	.69	.65	.75	.59
	Science CRI	.65	.45	.71	.57
	Science Theta	.70	.56	.74	.62
	Social Studies CRI	.64	.70	.69	.63
	Social Studies Theta	.66	.65	.72	.60

11; however, the variation in the performance difference is still larger in school districts with higher percentages of minorities.

The performance difference is also larger between African-American and white students in schools with above average percentages of minority students (See Table 11). Whites are always less likely to outscore blacks if they attend schools with below 3% minority populations. This variation in the performance difference is more pronounced in Grades 7 and 8, and again the variation in the difference at Grade 11 is much smaller.

An examination of the interaction between the racial mix of the school and the district shows several interesting patterns. One of these patterns is that the size of performance differences in reading and science between whites and African-Americans are smallest for students attending schools with low percentages of minorities in districts that have high percentages of minorities. Another pattern is that students in districts with below 3% minority enrollment and in schools with below 3% minority enrollment have the smallest performance differences in math and science at the 4th grade level and math and social studies at the 11th grade.

However the most intriguing pattern comprises students who attend school districts with less than 3% minorities but who are in schools with high minority percentages. At the lower grades (4 and 5) the differences between whites and African Americans are fairly stable. The 4th grade reading performance difference is smallest for these students as is the 5th grade social studies difference. This pattern changes in these schools at grades 7, 8 and 11. In 8th grade the math and social studies performance differences are the smallest compared to the other district and school pairs while the reading and science differences in 7th grade in these same schools are the largest. For example, in districts with below median minority enrollment but in schools with above median minority enrollment white students are likely to score higher than black students in science 80 to 87 percent of the time, depending on which measure is used. The same is true of the 5th and 8th grade social studies performance differences and the 8th grade mathematics performance difference. The 11th grade performance differences in these same district/school pairings are the highest performance differences of the four possible school and district pairings.

Summary of Exploratory Study 1

This study was a preliminary examination of the proposition that the geographic concentration by gender and race affected the size of the performance difference in schools and school districts. This was a cross-sectional study using data from only the 1997-98 academic year. As predicted, gender concentration did not appear to affect student performance on the KIRIS test substantively. However, the racial concentration of students in both schools and school districts does appear to have sizable effects on student performance.

Exploratory Study 2: Qualitative Examination of Schools with Low White/African-American Performance Differences

A preliminary draft of the current report, examining performance differences within Accountability Cycle 2 (1993-1996), began during the primary researcher's two-year tenure with KDE's Office of Curriculum, Assessment, and Accountability (OCAA).²⁹ In response to finding that a performance difference between Kentucky's white and African-American students was increasing over Accountability Cycle 2, this researcher was charged with seeking out and examining schools that had consistently small racial performance differences. This qualitative study attempted to identify characteristics of Kentucky schools and/or school policy that lead to the small differences in performances between each school's white and African-American students.

In this study, schools with the smallest performance differences over Accountability Cycle 2 were identified and six were chosen for site visits.³⁰ The selection process involved calculating the difference between white and African-American theta scores for each of Kentucky schools (n=1392) for each year of Accountability Cycle 2.³¹ In an attempt to ensure that the difference between white and African-American means was not being calculated on extremely small (and therefore unstable) white or African-American student samples, only schools that had at least 15 white and 15 African-American students in each accountability grade each year were included in the selection process. This selection criterion reduced the number of schools in the selection pool to 151.

Final selection for inclusion in this exploratory study was made in part on geographic diversity. Schools were chosen to maximize the number of different school districts included in the study. In addition, some schools were chosen specifically because they were middle and/or high schools, where a finding of consistently small performance differences across Cycle 2 was rare. Each of the six schools selected was in the top ten of schools with the smallest racial performance differences over Accountability Cycle 2.

²⁹The Office of Curriculum, Assessment, and Accountability no longer exists within KDE. During the 1998 restructuring of KDE, curriculum was divested from this office to become the Division of Curriculum Development within the Office of Academic and Professional Development. The remaining functions of the office (i.e., assessment implementation, issues concerning equating, research concerning assessment and accountability, and validation studies) are now housed in the Office of Assessment and Accountability or contracted to outside researchers.

³⁰A seventh school, in Fayette County, was also visited as a test of the semi-structured interview. While this school had a very small performance difference between its white and African-American students, it did not meet the requirement of having 15 white and 15 African-American students in its Accountability grades over the entire cycle.

³¹During Accountability Cycle 2, KDE gave all four major subtests in the three accountability grades (4, 8, and 11). Because KDE gave all four tests to the students at each grade, a composite measure could be constructed for each student. It was the white average composite theta and the African-American average composite theta that were used in the calculation of the smallest performance differences during the selection process.

To examine whether the selected schools were representative of Kentucky schools' performance, each school's accountability index for each year of Cycle 2 was compared with each year's state average for their grade levels assessed. Table 13 indicates that the schools selected tended to score somewhat below the state average for their grade levels assessed. However, at least two schools in the sample had higher scores than the state average across most years. Therefore, a hypothesis that only low scoring schools can have small performance differences is not generally supported. Half the selected schools exceeded their achievement goal for Accountability Cycle 2, suggesting that at the time of the study these were schools on the fast track to higher KIRIS scores.

Table 13. Grade Level, Accountability Indices, and Percentage of Goal Achieved over Cycle 2 for Schools Included in Qualitative Study (Average state index scores for each grade level are in parentheses.)

	Grade Level	Accountability Index 93	Accountability Index 94	Accountability Index 95	Accountability Index 96	% of Goal Achieved over Cycle 2 by school
School A	7-12	31.1 (36.2)	46.1 (42.6)	48.5 (44.9)	42.2 (42.7)	91.8%
School B	P-5	24.1 (35.7)	52.4 (40.9)	72.8 (46.2)	52.9 (45.0)	406.5%
School C	6-8	27.1 (37.4)	34.4 (41.8)	40.0 (44.9)	32.8 (41.9)	69.6%
School D	6-8	34.5 (37.4)	40.0 (41.8)	43.3 (44.9)	38.7 (41.9)	44.4%
School E	P-5	33.9 (35.7)	38.3 (40.9)	46.8 (46.2)	37.5 (45.0)	121.9%
School F	P-5	27.4 (35.7)	32.7 (40.9)	38.1 (46.2)	40.2 (45.0)	141.4%

Once the sample was selected, visitations were arranged at each school. KDE's OCAA member of the Regional Service Center (RSC) that served the targeted school made arrangements for the visit to each targeted school. The RSC staff member was invited to accompany the researcher on the school visit. It was believed that the inclusion of this person would ease potential distrust on the part of school personnel. Historically, a friendly fact-finding visit by OCAA had been exceedingly rare.³² However, the work schedules of most RSC personnel during the weeks before and immediately after KIRIS assessment administration prohibited most of them from accompanying the researcher.

³²One of the OCAA's branches was charged with surprise inspections during the KIRIS assessment to look for any school district or school violations in the administration of the KIRIS assessment. While researchers under contract with KDE often visited schools for research purposes, a visit by OCAA members for the purposes of research was very rare.

School visits were made in mid-April and late May 1997 in the period just before and after the KIRIS assessment testing window. Given the short length of time between the end of the KIRIS testing window and the beginning of summer vacation, there was no time to visit a random group of similar schools with larger performance differences to develop a baseline, nor was there time for an extended visit with any of the schools examined. Each school visit included interviews with the principal, the guidance counselor, and as many students and faculty as each of these daylong visits permitted.

While it might have been possible to force a rigid interview schedule on respondents, this exploratory study was designed to be as flexible and as unobtrusive as possible in hopes of limiting the amount of ‘staged performance.’ Sometimes the interviews occurred individually, but more often small groups of teachers and/or other personnel would be interviewed together. In two of the locations, the principals invited students to give their viewpoint on the reasons for the small performance difference.

In each school, the first question in the semi-structured interview asked the faculty and students for their ideas on why the performance differences between whites and African-American students were so small at their schools. After the respondents completed their answers to this question, a series of questions derived from previous research on achievement differences were asked.³³ Topic areas within the interview included: curriculum issues, instructional issues, professional development opportunities, leadership, interaction of faculty with parents and community, and the “inner workings” of Extended School Services and Family Resource Centers. Other areas mentioned by respondents were also explored during the interviews.

Because of the very exploratory nature of the qualitative study, the limited time for interviews and the small sample size, the results of this study must be considered very tentative until those aspects unique to schools with low performance difference differences can be separated from those aspects common to all Kentucky’s schools. It is recognized that school personnel could employ many different practices that might decrease the performance differences evident in their test scores. For the purposes of this report only those areas or practices discussed by school personnel at more than one school were considered.

School Vignettes

Before presenting the data gathered from the school visits, having some background information on the schools in the study is important. These short vignettes provide some context for understanding the similarities and differences between the populations each school served.

³³These questions can be found in Appendix J.

School A

The first school in the sample was a combined middle school and high school in a small, rural independent district in central Kentucky. This district served two small communities, each with around twenty-five hundred residents. While this district was the only totally rural school district in the sample, the major source of income for the area was not agriculture, but manufacturing. Only 4.5% of the school district's population was in households on farms (NCES 1994). In fact, the USDA's Economic Research Service labeled the county within which this district is found to be manufacturing dependent (i.e., manufacturing contributed a weighted annual average of 30 percent or more of total labor and proprietor income over the three years from 1987 to 1989.) (Cook and Mizer 1989). According to the guidance counselor at the school, the major employer for the school district was a Styrofoam cup manufacturing plant. Others residents often worked in tourism. This was most likely because the school district is close to a major park. Still these two economic bases, usually considered important for stable rural community development, had not prevented 23% of the total population (24.5% of the children) from living below the poverty line (NCES 1994).

School B

School B was an elementary school, which was part of a county school district in the Pennyroyal region. The county was part of a Metropolitan Statistical Area in western part of the state; however, about a third of the school district's population lived in rural areas (U.S. Census Bureau 1999; NCES 1994). The largest employer in the county was a U.S. military base, which provided almost half the earnings for county residents. At the time of the survey, the faculty claimed that the county had one of the lowest unemployment rates in the state. Again, however, a low unemployment rate did not always equal a good livelihood. According to the National Center for Educational Statistics (1994), 20% of the school district's population (and 30% of its children) was living under the poverty line. In this school, the faculty claimed the school to be 40% black with around 92% of the children on free and reduced price lunch. Before redistricting to include portions of the rural population, school officials stated that the school had been about 80% black. The faculty estimated the mobility rate of the students to be near 10%.

School C

School C was a middle school in the northern part of Lexington-Fayette County, the state's second largest city. Over the past ten years, Lexington had moved closer to a postindustrial economy with the largest and fastest growing sectors being service and government and the smallest and slowest sector being durable goods manufacturing (Bureau of Economic Analysis 1997). About 14% of the total population (18% of the children) were in poverty (NCES 1994). While only 21% of the children in Lexington were minority status in 1990, the principal estimated School C had a minority population of nearly 50%. Suggesting that poverty cut across race lines at School C, the principal estimated 75% of the students received free or reduced price school lunches.

School D

School D was a middle school in a county adjacent to one of the state's metropolitan areas. According to the Cook and Mizer (1989), this county's economy was dependent upon government employment, specifically on a nearby military base which had provided more than 25% of the county's income. The county was nearly 40 percent rural; however, only about 4.5% of the population resided on farms. This school district had a much lower poverty rate than the other sample schools. Thirteen percent of the population and 16% of the children fell below the poverty line (NCES 1994). According to the principal of the school, the school, like the county, had a decidedly military flavor. Students were most likely to be related to currently serving or retired military personnel. This school's service area was not a transient military population, but a more stationary, lifetime military neighborhood.

Common Characteristics of School E and F

Schools E and F were both elementary schools in the Jefferson County Public Schools (JCPS) system. Jefferson County was the state's most populous county. Like other metropolitan areas, Jefferson County had a strong service economy. It also had a strong manufacturing base although it had declined slightly over the previous ten years (Bureau of Economic Analysis 1997). The JCPS district included 662,544 persons; twenty-one percent of the total population were children. Minorities accounted for about 18.5% of the total population and one quarter of the child population (NCES 1994). Approximately, 13.5% of the total population and 20% of the children were living below the poverty line in 1989.

School E

School E, in Southern Louisville, served many children living in poverty. This school was the last elementary school that JCPS built in the city and was placed next to a large, integrated, housing project which still supplied the school most of its student body. According to the staff the school had a 50% minority enrollment. The school district transported (i.e., bused) few students to the school. According to the faculty, more than 90% of their students have been classified as "at risk." They also stated that the student body had a fairly high mobility rate among students as their families attempt to pull themselves out of poverty and leave the housing project.

School F

School F was an older elementary school near Churchill Downs in south central Louisville. The faculty stated that the school has around 35% minority enrollment. The white children that attend this school lived in the local neighborhood; the school district transported the African-American children in from the west end of Louisville. Still the faculty believed that all their children had nearly the same socioeconomic status. The counselor reported the mobility rate for the school's service area to be around 25%. However he also remarked that many

students returned to the school saying “if kids leave in September, they’ll most likely be back by February because this is the cheapest housing in town.”

Additional Characteristics of the School Sample

Each school system reported additional information to the state about school enrollment and participation in federal school nutrition programs. The National Center on Education Statistics also aggregated much of the 1990 census data to the school district level on its CD-ROM, the School District Data Book. Examining the information reported by the schools suggests that school personnel slightly overestimated the actual percentage of minorities within their school.

Table 14. Grade Level, Total Enrollment, Percentage of Minority Enrollment, Percentages of Free and Reduced Price Lunch Participants and Percentage of Urban Students included in the Qualitative Study

	Grade Level	Total Enrollment by school	% Minority in school	% of Students on Free Lunch in school	% of Students on Reduced Price Lunch in School	% of Urban Students in School District
School A	7-12	438	17.5%	27.4%	5.0%	0.0%
School B	P-5	387	35.5%	68.5%	12.9%	64.6%
School C	6-8	590	48.9%	54.7%	11.7%	97.0%
School D	6-8	758	22.4%	24.7%	13.3%	57.9%
School E	P-5	575	43.9%	92.9%	6.4%	96.3%
School F	P-5	635	26.7%	69.1%	11.7%	96.3%

As indicated in Table 14 most of the sample schools had more than 50% of the student body participating in the Federal school nutrition program. Free and reduced price lunch participation rates are often used as indicators of how poor the student body is. In order for a student to be eligible for free breakfast and lunch, the student’s family has to have an income no more than 120% of the federal poverty line. If one’s family income is between 120 and 185% of the federal poverty level, one qualifies for reduced price lunch (U.S. Department of Education 1996). School A, the combined middle and high school, had the lowest participation with just less than a third of its students participating in the program; however, KDE personnel have noted that six years of empirical data show that there is a substantial reduction in the use of federal school lunch programs from elementary to high school (Wetter 1999).

The last column in Table 14 provides some indication on the overall concentration of population within the school district. As can be seen, there is only one school district with no

urban students present. Every other school district in which a school was examined had more than 50% of its students coming from urban areas.

Preliminary Findings of Exploratory Study 2

The results of this exploratory study can be divided into factors that faculty/faculty and faculty/student activity can affect and interaction and factors that they cannot affect. When the faculty at each school was asked why their students showed such a small performance difference, the first factor mentioned in nearly every school is a characteristic of the student population that is external to the school. In each school, school personnel suggested that the small performance difference was a function of similarities in the socioeconomic status of whites and African-American students at their school.

Factors Beyond the Purview of the School Faculty

Student Social Class

Social class is a term used to refer to people who have similar levels of wealth and income (Macionis 1999). When sociologists and other researchers wish to measure social class, they often use socioeconomic status as an indicator. Socioeconomic status is derived through scaling of a person's education, occupation, and income in an attempt to assess an individual's social class, social status, and power, with a single measure (Keeves and Saha 1994; Frisby 1998). In this study social class and socioeconomic status will be used interchangeably.

In America, children have the social class of their parents or caretakers until they exit the educational system and enter the job market. Like geographic concentration (which happens to be a correlate of social class), socioeconomic status is an important concept in social science because it, too, is believed to affect individuals' experiences and life chances (Parker, Greer, and Zuckerman 1988; Wilson 1987; Zill, Moore, Smith, Stief, and Coiro 1991). The most recent decennial census data for Kentucky indicate that African-American students' families were more likely to have lower incomes than white students' families (NCES 1994). Thus, it is hypothesized that the effect of family's economic status on student performance becomes confounded with the effects of the race on student performance. Extending this argument, if the social class of whites and African-American students are equal then the performance differences between whites and African-Americans would disappear, or at least be reduced.

Faculty members at each school in the study believed, that the African-American and the white students performed similarly on the KIRIS assessments because most students' families in the school had similar social class. In most of cases, the faculty perceived that their students' shared background was lower class. Looking back at Table 14, only one school had less than 50% of its students on free and reduced lunch; in the rest of the cases the proportion of student using the federal program was much higher [between 67 and 89%]. In five of the six schools, at least one faculty member stated that the group of students with the lowest socioeconomic status was actually lower class whites, which suggested that African-American students were

sandwiched between middle and upper class whites on one side and whites in extreme poverty on the other.

A faculty member at one school went on to suggest that class affected performance on the KIRIS assessment because it limited the types of experiences that a child would have. He stated that kids with a variety of experiences do better on the assessment. If the families do not travel much, the children have very little experience outside their neighborhoods. In the schools in this study the students were economically similar and therefore both whites and blacks could be expected to have the same set of limited experiences. It might be that the similarity of experiences makes it easier to keep students together in the learning process and thereby keep their performance difference smaller. As an example, this faculty member cited one classic case in which class differences had affected a KIRIS question. In this question, students were asked to write about what three places in Kentucky they would take a foreigner or an alien to see. The teacher suggested that the typical response in the school in which he or she taught would be grandma's house or the local mall rather than Natural Bridge. Moreover, a principal at another school argued that field trip experiences were the only real way to expose students to the fine arts upon which the KIRIS assessment tests.³⁴ Thus, in their attempts to provide more varied cultural experiences for their students, at least two of the schools in this study have begun holding more field trips.

Three other variables that were mentioned earlier in the vignettes are considered important when considering the effect of socioeconomic status on performance differences. Redistricting, transporting of students, and consolidated high schools may throw lower class students in with students of different class backgrounds and educational experiences. Calabrese and Poe (1991) have suggested that class differences may lead to frustration and alienation, especially among minorities. Again, it may be that a differential dropout rate affects performance differences. When alienated students are given a chance to leave the system, a great number do so (Calabrese, 1991, Trusty and Dickey, 1993).

Finally, at least one faculty member in one metropolitan school posed the idea that the small performance difference was actually an artifact arising from magnet schools attracting their school's best students. In his view, the "cream of the crop" within the school's neighborhood was more likely to transfer to a magnet school, by that level student achievement in school they would have normally attended. Of course, such an effect could only occur in school districts where private schools, magnet schools or another type of school choice program were in effect that could draw students away from the public school system, and therefore cannot explain the lack of difference at all six of these schools.

³⁴Although this exploratory study was concerned primarily with minimal student performance differences between whites and African-American students, respondents typically view the KIRIS assessment holistically, not as a series of subtests. The principal's comment in this case refers to the fact that the assessment tests student performance in the areas of arts and humanities. At the time this study was done, many schools in the state were attempting to improve their students' knowledge of the arts and humanities because student scores statewide on this subtest had been low the previous year.

Proportion of Black Students in the Schools

A faculty member at one school stated that the racially balanced nature of the classroom was also important in keeping the performance difference between whites and African-Americans small. Previous social theory and the current data support this hypothesis. In all six schools visited, the percentage of blacks enrolled was at least 6 to 16 times greater than the state's median percentage enrollment (3%). While this may seem trivial, sociologists have found that shifts in the proportion of minorities within a group often transform group interaction (Kanter 1977).

When the proportion of minority members is very small, Kanter (1977) predicts that they will face three problems: performance pressure, boundary heightening, and role entrapment. Under performance pressure, members of minority groups always feel that others are scrutinizing them (Jackson et al. 1996). They also feel pressure to act as a representative of their social group. Yet no matter how well they perform, minority group members feel that their minority status always has more affect on their performance than their actual achievement does. Often feel that they have to do more to be noticed. The second problem faced by minority group members is boundary heightening. This is the exaggeration of in-group commonness and out-group difference. Minority group members are repeatedly reminded of their differences through jokes, interruptions, exclusion, and various 'loyalty tests'. Last, minority group members face typecasting by dominants. This role entrapment maps minority group members' actions onto existing cultural stereotypes.

For minority group members there appear to be some common solutions. One solution to these problems is to over achieve and be accused of selling out to the majority. Another is to avoid calling attention to oneself and, thereby, suffer a lack of recognition and an increase in alienation and social isolation. However, these problems described by Kanter decrease as the proportion of minorities increases, allowing minority members to form alliances and have more input on the culture of the group. The more equal the proportions, the less members of the minority group attract attention based on their race and therefore the less minority membership stigmatizes or stereotypes the individuals (Kanter 1977). The faculty member who discussed classroom balance reported that a balanced classroom promoted self-esteem, suggesting that each individual in the classroom feels as equal as everyone else in the room. In these cases, the less performance pressure placed on the African-American students, the better able they are to perform academically.

Proportion of Minority Faculty

Faculty and staff at each school made sure to point to the number of black role models that were accessible to their students. Research suggests that the presence of these role models should also reduce the performance differences between black and white (Evans 1992). Some middle school faculty members were concerned that they are the only positive role models that their students have whether black or white. Some schools did have more blacks in the classroom and in the main office than others. On the high end, one elementary school reported that its

guidance counselor and 20 to 25% of its faculty were African-American. Other schools studied did not have as many African-American instructors. In one elementary school, the principal reported that he or she tried to make sure that each grade had at least one quality minority teacher. The staff also pointed out that African-Americans worked in very visible positions within the school, such a cafeteria manager and janitorial service. Two of the schools attempted to supplement the impact of visible African-American role models by bringing in motivational speakers to reinforce the idea that students can achieve their dreams if they work hard.

However, when speaking with several African-American teachers in the schools studied, a disgruntled feeling was noted. This feeling is due in part to the fact that African-American teachers believe that they are seen as the primary role models African-American students. In addition, some African-American teachers also perceived that the school administration was asking them to deal with more of the discipline cases. Based on these understandings and perception, these teachers believe that they should be getting more pay.

Factors within the purview of the faculty

While socioeconomic status and the proportion of minorities in the school are two variables which may be important in determining why the sample schools have low performance differences, there tends to be very little that school faculty can do to change these factors. There are however several factors that faculty can and do change in their attempts to increase student performance.

Curriculum

At its minimum, a curriculum is a set of planned learning experiences for a particular area of study (SSTA 1999). Curriculum content was different at each school studied. While curriculum content may be an important factor in determining why students at these schools had small performance differences, a true examination of content cannot occur in a one day site visit and was not attempted. When the faculty members at each school were asked how they connected the issues of race to the curriculum, two schools' personnel suggested that knowledge of cultural diversity was important. For example, one school held a "Christmas around the world" party to illustrate how different ethnic groups celebrated the holiday. The only other comment on this topic came from a middle school history teacher, who noted that students in the school became uncomfortable during classroom discussions of slavery.

Curriculum Alignment and School Transformation. While the curriculum content and structure in the sample are unknown in this study, the effect of faculty commitment to two of the initiatives of the Kentucky Education Reform Act (KERA), curriculum alignment and school transformation planning in these schools can be examined. One major emphasis of KERA was on community and faculty participation and empowerment (Keith 1996). In this process KDE gives each school feedback through KIRIS concerning how its students perform. In addition, KDE has developed core content guidelines in each area assessed that identify what is essential for all students to know about that topic by grade (KDE 1999a). The state, through KDE, then

charges the community and the faculty to change their schools to improve student performance. Community members work through the School-Based Decision Making councils; faculty work on school transformation plans to develop the overall plan for improvement then work together, usually in teams, to align their curriculum so that it better resembles the core content. In each school examined in this exploratory study, faculty members drafted and approved their school transformation plans. Faculty members felt that this gave them a voice in exactly how they would carry out KERA at their school. The buy-in on curriculum alignment and school transformation suggests larger trends in these schools and occasionally their school districts. For example, the faculty at one school studied had been involved in coordinated planning to create writing and math portfolios from preschool through 12th grade. Several the variables that relate to curriculum alignment and school transformation planning also concern leadership and will be discussed further in that section.

Instruction

Curriculum, as noted above, is a set of planned learning experiences; instruction is the process by which students are guided through the curriculum. Good instruction is therefore tied to making sure that the students get the most out of the curriculum. The most consistent finding in this study across schools regarding instruction was that faculty in at least three of the schools, argued that setting high expectations and challenging children to learn was an important part of good instruction. These teachers stated that they expected the same from all students, while they attempted to allow for individuality, and in the end, suggested teachers needed to praise their students for their accomplishments.

At one school a teacher stated that according to Vygotsky, kids can learn together when they are at the (achievement) level they should be at, one level below what they should be or one level above where they should be. At her school, given the student backgrounds the teacher felt that this was pretty much the case. However, when kids get bused, they often get put in schools at which they are beyond one level's difference from the norm. The teacher and her colleagues stated that they are forced to go slower and be more repetitive in their instruction than other teachers at the same grade level in other schools. One teacher gave the example that she had a child in another school in the district, who was currently in the same grade level that she taught, but that her child was well above what she was teaching in her class. The teachers stated that if they switched schools that they would have to change the tempo, pace, and language of their instructional techniques. The idea that their students were not keeping pace with other students at the same grade level concerned these teachers. They noted that since their school tended to have an extremely high mobility rate, students would be entering the school that they were transferring to with a pre-existing achievement deficit because of the instructional pacing. The teachers went on to note that many students who transfer return to the school after only a few months. At that point they again would be behind in the instructional pacing in their original school. Therefore, the suggestion that minority students are behind at a particular school may not be a result of discrimination, but may instead depend on instructional pacing at specific schools. However, they went on to point out that being behind is not inherently discriminatory.

Instructional Teams. Four of the schools in this exploratory study were large, having student enrollments greater than 500. However, such large enrollments put certain types of students at a disadvantage. Current research on the effects of school size suggests that small schools provide an achievement advantage for students with lower socioeconomic statuses (Howley 1996; Plecki 1991).³⁵ Thus the four large schools studied were engaged in simulating small schools through "schools-within-schools" programs (Raywid 1996). In these programs, the administration divides the schools' students and faculty into two or more separate schools within the same building. Each school runs as if it were a separate but contiguous school under one central administration. There were two variations in the schools-within-schools programs that administrators had set up these programs within the sample schools. Two schools ran teams, but did nothing else differently. Two other schools were engaged in student-teacher progression programs. In these programs, teachers stay with the same group of students through several school years.

Faculty involved in these schools-within-schools programs stated that they had strong parent involvement, which meant that they stated they knew more about their students and were better acquainted with the students' families and family situations. This was because the faculty got to know the students better and had more chances to meet parents. While this strategy appeared to get parents more involved and may have helped kids achieve while they were in middle school, the teachers were concerned that once students left their school, and went to a larger traditional high school where no student-teacher progression program existed, the students might suffer both socially and academically. These teachers feared that the teacher-student hands-off approach of the traditional school setting would affect their academic performance.

In three of the four schools in the sample, the team memberships were very stable although there was concern at one school that the principal would soon be trading team members to maximize student achievement, sacrificing teams whose personalities were complementary and who worked well together for teams that could probably boost academic performance higher, but would include conflicting personalities. (Education Trust 1999).

Discipline Systems. The teachers at three schools mentioned that their swift and objective discipline systems that appear to give students clear guidelines to follow thus focusing students on academics. At another school, a teacher described its efforts to include peer mediation of discipline problems trying to limit behavioral problems.

Instructional Technology. Another very dynamic aspect of instruction in schools has been the increase in amount and type of technology available for instruction. The schools in the sample had integrated technology into instruction very early. Several schools had state of the art computer rooms and/or other technological innovations within their buildings. For example, one middle school in the study was one of few schools in the state to have a Kentucky Tele-Linking

³⁵Other investigators conclude that no school should have more than 400 to 500 students because affect and social benefits are not as great as school populations continue past 500 students (Cotton 1996).

Network conference room. This room allowed the students access to some educational experiences that were unavailable to students at other schools. For instance, the school's students had been involved in a video conference with Nobel prize winners during the week before the school visit. Another example comes from an elementary school which had installed a new student computer lab, put a computer in every classroom, and had enough laptops so that only two students had to share a machine.

Other Instructional Resources.³⁶ In contrast to the amount that some schools spent on new technological innovations, several teachers reported frustration with the amount of their own time and money that they were forced to spend to provide quality instruction. Teachers from several schools commented on how much time they spent preparing for classes. In addition, other faculty pointed out how much money they spent from their own pocket for the school materials necessary for instruction. This is particularly the case, according to faculty, when developing materials to teach new concepts. For example, a science teacher at one elementary school reported spending nearly \$500 for materials for science experiments. A teacher at another school stated that because she was expected to bring the students closer to her while she taught, she bought a rug for them to sit on. Since she bought the rug, the school now refuses to clean it. Finally, teachers said they do not have enough storage space to organize and keep materials that they have already used or are preparing to use. In one classroom I visited, the teacher was using half the space in the room for storage of her course materials.

Professional Development

Professional development is “a continuous process of [faculty] improvement to promote high standards of academic achievement and responsible citizenship for all students” (MSBE 1997). The onset of KERA changed the nature of professional development in Kentucky. According to the faculty in the schools studied, most professional development activity was occurring internally. As one teacher succinctly stated, the benefit to such a situation was that everyone in the school receives the same message; however, the downside was that their school's faculty could become isolated from other schools' faculties. The only teachers that could attend professional development activities away from their school were individuals involved with student organizations that run on their own funding (e.g., FHA). One school attempted to compensate for the potential isolation by hosting the county summer school at their facility. This allowed their faculty to mingle with faculty from other parts of the district and trade classroom ideas.

Leadership

Under KERA, the definition of leadership has changed from being the process by which one individual coordinates group activity to mean the establishment of “a learning community

³⁶The resources of time, space, and money are discussed here. It is recognized that the faculty have more control over some of these resources than others.

through facilitation, collaboration, coaching and shared accountability” (Michener, Delamater, and Schwartz 1990; KDE 1999c). Leadership issues were most often discussed by staff in the sample schools by talking about the school transformation planning and curriculum alignment. While the faculty at some schools noted that they did not always agree on the appropriate course of action to take to transform their school and/or align their curriculum, they suggested that they had the skills and the commitment to come to a compromise and move forward.

While group decision making and planning may empower the faculty, this empowerment comes at a price. Unless planning meetings can be kept brief, faculty members begin spending larger amounts of their time in planning. The faculty at the two middle schools studied noted that beside their individual planning time, most teachers also had group planning time together during the day, so that they could decide logistics among their group, grade and/or team. In one elementary school, the reading teachers evaluated student performance as a group to decide the teaching method to employ (phonics vs. whole word instruction). This faculty empowerment seemed to come at the expense of the principals’ traditional leadership role. At least one principal expressed dismay that he could not “ram stuff down the teachers’ throats” because sometimes they got so involved in these projects that he could not get them involved in other things.

Networking: Parent Involvement, Community Involvement, School-to-School Ties

There were faculty members in each school who said that parents were not very involved in the school’s efforts to educate their children. Other faculty pointed out that the level of parent involvement with the school differs based on grades within the school. Usually, this meant that grade schools had much parent involvement, but day-to-day involvement tended to slack off in middle school with parent energies getting channeled into PTSA and sports and band boosters.

Personnel at most of the sample schools suggested that they went great lengths to try to increase parental involvement. According to reports by the faculty at five of the six schools, these initiatives failed to increase involvement substantially. Elementary school faculty appeared to attribute this lack of volunteerism to the parents’ social capital. For example, one faculty member acknowledged that most of the parents near his school had to work and those that did not work were probably too insecure in their own abilities to feel that they could help their children learn. At this school, parent involvement has come to mean making sure that parents urge their kids to do their homework. The faculty member suggested that the school also distributed strategies for dealing with behavioral problems.

Other types of networking appear to be less frequent. Only one school reported attempting to facilitate the development of neighborhood ties between the school’s students and prominent community figures. This school offered special lunches (e.g., grandparent days, parent days, media days, minister days, civic leader day) to which students are allowed to ask their extended families and/or other neighborhood individuals. School to school networking does not appear to occur outside professional development efforts. Faculty in only one of the sample

schools reported doing a presentation at another schools the district. This faculty member of School F reported giving talks at School E.

Extended School Services and Family Resource Centers

Each school included in the sample had some form of Extended School Services (ESS) they offered to their students. However, each school's program differed from the program at the other sample schools. While it is conceivable that extended school services might help keep performance differences low, there were so few similarities in the school's programs that any consistent effect would be hard to explore in this study. Some school faculty reported that their school ran summer programs while faculty at other schools reported that their school did not. Some faculty reported that their school programs accepted every child who wished to stay after school, faculty at other schools reported that their ESS programs have strict entrance criteria. Most of the schools offered aid in remedial work. In two of the schools at least part of the ESS program staff worked with students on their strategies for answering KIRIS style test items.

Besides their ESS programs, members of the faculty of at least three of the six schools studied suggested that their Family Resource Center had really helped improve students' lives and therefore their classroom performance. Still, the exact offerings of each Family Resource center are unknown. Future research needs to examine the types of programs offered more thoroughly and their effects on students' achievement.

Summary of Exploratory Study 2

This exploratory study examined six schools looking for common factors that could explain the small racial performance difference at each school over Accountability Cycle. To carry out this exploratory study, semi-structured interviews were conducted at each sample school. The interview schedule consisted of questions that current research literature suggested would affect student performance. The faculty, staff, and students were eager to describe and discuss their schools in a general sense; however, school personnel perceived few factors that they could point to for an explanation about why their racial performance difference was so small. When analyzing the data from all six schools, the faculties of every school commonly reported only one variable that they believed affected scores: socioeconomic status. Other factors such as the high proportion of black students in the school, the provision of Extended School Services, and/or the practice of providing challenging instruction for all students were reported by staff from most schools; however, the faculty and staff at the sample schools did not perceive these factors as reasons that the racial performance difference was small. In the literature many other factors have been considered important predictors of racial performance differences.³⁷ The effects of these many types of variables on racial performance differences at the schools included in this exploratory study were not clear, and in most cases, untestable given the exploratory nature and the time constraints of the study.

³⁷While the references for this report are not considered to be exhaustive, all the following articles note various factors often considered related to differences in racial academic performance: Alexander 1996, Banks, McGee and Banks 1995, Boone 1997, Chandler 1997, Clark 1991, Collins-Eaglin 1993, Evans 1992, Fordham 1988, Frisby 1998, Greenwood Terry, and Utley 1993, Jencks 1998, Kennedy 1995, Nweze 1993, Page and Rosenthal 1990, Roscigno, Rosenthal 1973, Rowley 1996, Trusty 1993

Exploratory Study 3: Quantitative Examination of Race, Gender and Socioeconomic Status

As noted in the summary of Exploratory Study 2, there was a widespread perception among staff and faculty in the schools under study that when students' social class was similar the performance differences between whites and African-Americans either became smaller or disappeared. While such a hypothesis was straightforward and potentially testable, getting data on students' social class has been a difficult proposition. Exploratory Study 3 was designed as a further exploration of the possible effects of social class on student performance using data that were currently available from KDE. This quantitative study focused on two regression models. The first model examined only the relationships among race, class, and gender and their possible effects on student performance. In the second model, seven additional variables from the KIRIS assessment student questionnaire were added. These variables serve as indicators of some concepts that have been discussed in earlier studies.

To test the hypothesis that equivalent socioeconomic status reduces or eliminates racial performance differences, data beyond that gathered as a part of the KIRIS assessment is necessary. Because KDE does not collect or have access to information on students' family and/or household incomes, the only data that KDE has on students' socioeconomic status at the student level comes from information collected concerning student participation in the Federal Title I program. Thus, a quick overview of Title I's role in Kentucky's schools is helpful.

Title I, Part A provides monies to local education agencies to assist disadvantaged children in learning. There are two different assistance programs within Title I, Part A. Schools with enrollments that are made up of at least 50% low-income students (i.e., students on free and reduced price lunch) are eligible to plan a schoolwide program. Here the Title I program pays for a schoolwide program that aids the entire student body. Schools ineligible for schoolwide program status may still receive funds from Title I; however, these schools must use any Title I funds received to provide services only to those children which the school identifies to be in greatest need. While school personnel apply for schoolwide program monies based on the economic disadvantage of the students, targeted assisted program monies may also be used to aid students disadvantaged by disability status, migrant status, or limited English proficiency status (KDE 1999b).

In 1997-98 KDE included an item on the "face sheet" of the KIRIS assessment that asked students attending targeted assisted school to identify whether the school was targeting them for Title I assistance.³⁸ To construct a measure of socioeconomic status, information concerning the

³⁸Prior to 1997-98 KDE had collected no information on social class at the individual student level.

school's Title I status and poverty level was combined with the information from the student assessment to develop a measure of Title I status to serve as an indicator of social class. Then each student's data was marked as participating or not participating in Title I.

Title I participation was coded based on three variables. Two of these variables came from the KDE's Title I data file and the other was reported in the KIRIS assessment's Student Control file. The three different ways that a student was marked as participating in Title I were: 1) If the Title I file showed that the school the student attended participated in Title I and had adopted a schoolwide program, then all the students in that school were marked as Title I program participants. 2) If the Title I file showed that the school participated in Title I and had adopted a Targeted Assisted Program, then the student records were checked to see whether the student had marked that he or she was a Title I targeted- assisted program participant. 3) Students in such schools, who had marked their face sheets appropriately, were also listed as Title I participants.

Several inaccuracies may occur in the Title I variable. The 1997-98 school year was the first year in which the self-reported Title I information participation was asked on the KIRIS assessment form. Schools have been slow to adopt the new question. Many schools, evidently, did not ask their students answer this question. This underreporting leads to inaccuracies because students are not flagged as Title I participants although they do participate in the program. An additional error can occur in the upper grades because disadvantaged students in these grades often do not participate in Title I (Wetter 1999). Even if they do participate, these older students are less likely to mark themselves on the form as participating. A third inaccuracy occurs because some students (perhaps as many as 50%) who attend schools that participate in Title I schoolwide programs will be marked as Title I although they are not economically disadvantaged.

In an attempt to limit these inaccuracies for Study 3, three selection criteria were developed to select only those student records with a higher probability of accuracy. The first criterion was that students were included if their schools did not participate in Title I. Approximately 44% of Kentucky's schools did not participate according to the Title I data file. The second criterion examined each schools' poverty level as reported in the Title I data file. Students in schools with schoolwide Title I programs were only included if the school's poverty level, as reported to KDE, was above 74%.³⁹ Using these criteria decreased the number of students being misreported as economically disadvantaged when they were not. The third criterion was used to decrease the error due to underreporting. The percentage of students targeted by the school was compared with the reported percentage of students in poverty at that school. If these percentages were within five points of one another, the students attending those schools were included. This not only eliminated schools that did not have their students mark their targeted assisted status but also ensured that most of the schools' students were receiving

³⁹This criterion was developed by researchers in the KDE's Office of Assessment and Accountability. The 74% poverty level was chosen by the researcher because it is the lower bound of the highest category reported in Title I documents.

targeted assistance due to economic disadvantage rather than another factor. Students had to meet one of these three criteria in order to be included in the study. In Table 15 are the percentages of students at each grade level who were included in the sample for Exploratory Study 3. The percentage of students at each grade level included in the sample increases at the higher grades. The increase in percentage of students included most likely occurs because fewer secondary schools than primary schools participate in free or reduced lunch programs (Wetter 1999).

Table 15. Percentage of Students Included in the Sample for Exploratory Study 3

Grade Level	Percentage of Students for Which Data Can Be Obtained
4	36.1%
5	36.6%
7	47.8%
8	47.6%
11	85.5%

Besides race, gender, and Title I participation, seven other independent variables were included in this study.⁴⁰ All are indicators of variables often seen as important in the academic performance of students. Three of the variables were alternative ways of measuring student geographic mobility. These three were length of residence in Kentucky, length of attendance at their current school, and number of times the student had moved in the past two years which required a change to another school. The fourth variable was number of days of absence in the last month before the assessment. The fifth variable was presence in student homes of a language other than English. The sixth variable was number of books in the home. The seventh variable was a self-reported overall grade.

In Tables 16 through 27 the standardized regression coefficients and the coefficient of determination for student theta (dependent variable 1) and CRI scores (dependent variable 2) are presented by subject area and grade, regressed on race, gender, and Title I participation. In Model 2 the seven other independent variables are added.

Generally the amount of variance in the dependent variables explained by only race, gender, and social class is low. Looking at the R^2 statistics for Model 1 in the tables, we see that the most variance explained by these three independent variables is 13% for 5th grade mathematics. More often the R^2 for Model 1 is between 7 and 10% in the 4th, 5th, 7th and 8th

⁴⁰These variables were taken from the student questionnaire that KDE administers as part of the KIRIS assessment. Each student questionnaire contains 50 questions to be answered by the students. Most of the questions vary depending on which form of the test the student is taking. The seven questions included here are the only questions that every student questionnaire had in common. The exact wording of these questions can be found in Appendix K.

grades. In 11th grade these three variables account for much less of the variance in the subject area assessments.

In spite of the low coefficients of determination, there are statistically significant independent effects of race, Title I status and gender on all subject areas at all grades. There are two exceptions. In the 4th grade gender is not a significant predictor of how one scores on the science performance assessment. The second exception is in 11th grade where Title I participation is not a significant predictor of either science or social studies scores. In the 11th grade, Title I participation does little for the prediction of the dependent variable. This finding may occur because few 11th graders are Title I participants. However, every time student social class (as indicated by Title I status) is taken into account there remains a significant independent effect regarding race, reducing support for the stated hypothesis. Thus the sample school teachers' perceptions that social class rather than race explains the small performance differences between black and white students are not supported by this analysis.

The analysis in Model 2 indicates that the seven additional variables from the KIRIS student questionnaire increases the amount of variance explained by the independent variables, but does little to change the relationships found in Model 1. The ten variables included in Model 2 explain nearly 20 to 25% of the variance in the theta scores and between 10% and 20% of the variance in the CRI scores. The standardized regression coefficient, indicate that students' perception of their grades appears to have the strongest predictive value. Students who report that their overall grades are poor or below average appear to do much worse on the KIRIS test, when the other variables in the model are controlled. Still, even when these new variables are added, with the exception of 4th grade, there continues to be significant, independent effects based on race and gender.

Only one of the three mobility variables--number of times the student has moved--included in Model 2 is significantly related to the performance measures. Neither length of residence in Kentucky, nor length of time in the school appears to affect student performance, once the other variables are held constant. The number of times the student had moved in the past two years which required a change to another school has a different relationship with performance at 4th grade than it has with performance at the other grades. In 4th grade, this variable has a negative effect on students' scores controlling for the other variables in the model. It may be hypothesized that disruption and stress occur in moving to a new school so that the more times a student changes schools, the worse he or she performs on the performance assessment. In every other grade we observe the reverse. Moves that require changing the school attended appear to have a positive relationship with student scores on the assessment.

The number of books in the home displays a pattern similar to that the number of moves variable. At the 4th grade, students whose families have more books do better in reading and science. This finding might suggest that more books in the home demonstrates a supportive reading environment and that students in such an environment do better on the KIRIS assessment. Other researchers have used this variable as an indicator of social class (Frisby 1999). As such, it would mean that being of higher social class has a positive effect on student

Table 16. Grade 4 Reading Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Reading Theta		Reading CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.28***	-.21***	-.24***	-.17***
African-American	-.14***	-.09***	-.11***	-.07***
Female	.14***	.13***	.13***	.12***
Length of Ky. Residence		-.02**		-.02*
Length of time at school		-.02*		-.02
Number of Absences		-.09***		-.07***
Mobility		-.12***		-.10***
Number of Books		.14***		.12***
ESL Home		-.05***		-.04***
Bad Grades		-.19***		-.16***
R ²	.12	.22	.09	.16
F	709.2***	405.3***	493.4***	276.6***
N	15191	14581	15191	14581

Table 17. Grade 7 Reading Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Reading Theta		Reading CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.17***	-.14***	-.12***	-.10***
African-American	-.16***	-.11***	-.11***	-.07***
Female	.25***	.20***	.18***	.15***
Length of Ky. Residence		-.01		.00
Length of time at school		-.01		.00
Absences		-.07***		-.04***
Mobility		.14***		.10***
Number of Books		-.10***		-.06***
ESL Home		-.04***		-.03***
Bad Grades		-.28***		-.22***
R ²	.12	.27	.06	.14
F	750.8***	594.7***	365.6***	271.6***
N	17270	16477	17270	16477

Table 18. Grade 11 Reading Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Reading Theta		Reading CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.02 ^{***}	-.02 ^{**}	-.02 ^{***}	-.02 ^{***}
African-American	-.10 ^{***}	-.05 ^{***}	-.08 ^{***}	-.04 ^{***}
Female	.28 ^{***}	.23 ^{***}	.24 ^{***}	.20 ^{***}
Length of Ky. Residence		-.01		-.01
Length of time at school		.03 ^{***}		.03 ^{***}
Absences		-.11 ^{***}		-.09 ^{***}
Mobility		.16 ^{***}		.14 ^{***}
Number of Books		-.06 ^{***}		-.04 ^{***}
ESL Home		-.03 ^{***}		-.03 ^{***}
Bad Grades		-.25 ^{***}		-.24 ^{***}
R ²	.08	.23	.07	.18
F	778.1 ^{***}	715.7 ^{***}	591.7 ^{***}	528.7 ^{***}
N	25285	24366	25285	24366

Table 19. Grade 5 Mathematics Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Mathematics Theta		Mathematics CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.29***	-.21***	-.27***	-.21***
African-American	-.21***	-.15***	-.17***	-.12***
Female	.05***	.01	.05***	.01
Length of Ky. Residence		-.02		-.01
Length of time at school		.02*		.02**
Absences		-.09***		-.07***
Mobility		.19***		.16***
Number of Books		-.09***		-.08***
ESL Home		-.03***		-.02*
Bad Grades		-.25***		-.23***
R ²	.13	.28	.11	.23
F	620.7***	458.7***	492.9***	340.2***
N	12165	11728	12165	11728

Table 20. Grade 8 Mathematics Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Mathematics Theta		Mathematics CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.16***	-.13***	-.14***	-.11***
African-American	-.20***	-.14***	-.17***	-.12***
Female	.06***	.01	.04***	-.01
Length of Ky. Residence		-.01		-.01
Length of time at school		.00		.00
Absences		-.09***		-.07***
Mobility		.18***		.16***
Number of Books		-.10***		-.07***
ESL Home		-.03***		-.02**
Bad Grades		-.27***		-.28***
R ²	.07	.23	.05	.20
F	419.6***	482***	285.8***	396.1***
N	16726	16139	16726	16139

Table 21. Grade 11 Mathematics Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Mathematics Theta		Mathematics CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.01 [*]	-.01	-.02 ^{**}	-.01 [*]
African-American	-.13 ^{***}	-.09 ^{***}	-.12 ^{***}	-.08 ^{***}
Female	.04 ^{***}	-.01 ^{**}	.02 ^{**}	-.03 ^{***}
Length of Ky. Residence		-.03 ^{***}		-.03 ^{***}
Length of time at school		.03 ^{***}		.03 ^{***}
Absences		-.12 ^{***}		-.10 ^{***}
Mobility		.17 ^{***}		.15 ^{***}
Number of Books		-.06 ^{***}		-.04 ^{***}
ESL Home		-.03 ^{***}		-.02 ^{**}
Bad Grades		-.31 ^{***}		-.31 ^{***}
R ²	.02	.21	.01	.18
F	166.9 ^{***}	643.3 ^{***}	123.8 ^{***}	541.8 ^{***}
N	25285	24366	25285	24366

Table 22. Grade 4 Science Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Science Theta		Science CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.23***	-.15***	-.19***	-.13***
African-American	-.21***	-.15***	-.18***	-.13***
Female	.01	-.01	.00	-.01
Length of Ky. Residence		-.03***		-.02**
Length of time at school		.00		.00
Absences		-.09***		-.08***
Mobility		-.11***		-.08***
Number of Books		.15***		.12***
ESL Home		-.04***		-.03***
Bad Grades		-.19***		-.15***
R ²	.10	.20	.07	.13
F	579.5***	358.2***	373.8***	216.7***
N	15191	14581	15191	14581

Table 23. Grade 7 Science Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Science Theta		Science CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.18***	-.15***	-.16***	-.13***
African-American	-.24***	-.18***	-.19***	-.15***
Female	.08***	.03***	.07***	.02***
Length of Ky. Residence		-.02**		-.01
Length of time at school		.01		.00
Absences		-.08***		-.05***
Mobility		.14***		.12***
Number of Books		-.10***		-.07***
ESL Home		-.03***		-.02**
Bad Grades		-.28***		-.26***
R ²	.10	.25	.06	.18
F	605.8***	539.5***	398.7***	354.7***
N	17270	16477	17270	16477

Table 24. Grade 11 Science Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Science Theta		Science CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	.00	.00	-.01	-.01
African-American	-.17***	-.13***	-.12***	-.08***
Female	-.03***	-.08***	-.03***	-.06***
Length of Ky. Residence		-.02**		-.02**
Length of time at school		.02*		.01*
Absences		-.11***		-.08***
Mobility		.18***		.12***
Number of Books		-.07***		-.05***
ESL Home		-.03***		-.02***
Bad Grades		-.23***		-.16***
R ²	.03	.17	.01	.08
F	258.3***	495.9***	121.6***	216.4***
N	25285	24366	25285	24366

Table 25. Grade 5 Social Studies Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Social Studies Theta		Social Studies CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.24***	-.17***	-.21***	-.15***
African-American	-.17***	-.11***	-.14***	-.09***
Female	.12***	.09***	.11***	.08***
Length of Ky. Residence		.00		-.01
Length of time at school		.01		.01
Absences		-.08***		-.07***
Mobility		.17***		.15***
Number of Books		-.09***		-.07***
ESL Home		-.05***		-.03***
Bad Grades		-.23***		-.21***
R ²	.10	.23	.08	.18
F	455.6***	353.7***	333.8***	254.1***
N	12165	11728	12165	11728

Table 26. Grade 8 Social Studies Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Social Studies Theta		Social Studies CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.18***	-.14***	-.14***	-.12***
African-American	-.19***	-.13***	-.14***	-.10***
Female	.09***	.04***	.08***	.04***
Length of Ky. Residence		-.03***		-.03***
Length of time at school		.02**		.02**
Absences		-.09***		-.07***
Mobility		.18***		.16***
Number of Books		-.11***		-.08***
ESL Home		-.03***		-.02**
Bad Grades		-.28***		-.26***
R ²	.08	.25	.05	.19
F	453.3***	536.2***	291.8***	368.1***
N	16726	16139	16726	16139

Table 27. Grade 11 Social Studies Theta and CRI Scores Regressed on Title I Participation, Race, Gender, and Other Independent Variables

	Social Studies Theta		Social Studies CRI Score	
	Model 1	Model 2	Model 1	Model 2
Title I Participant	-.01	.00	-.01	-.01***
African-American	-.12***	-.07***	-.10***	-.06***
Female	.18***	.14***	.16***	.12
Length of Ky. Residence		-.01		-.01**
Length of time at school		.03***		.03
Absences		-.11***		-.10***
Mobility		.16***		.15***
Number of Books		-.07***		-.04***
ESL Home		-.03***		-.03***
Bad Grades		-.27***		-.27***
R ²	.05	.21	.04	.18
F	417.3***	656.6***	310.6***	517.5***
N	25285	24366	25285	24365

performance even after Title I status and the other variables are controlled. However, in the later grades having more books around the house appears to have a very small negative effect on student scores. Perhaps this negative effect relates to the type of books available and the potential for leisure reading to be more exciting than schoolwork.

The last two variables—number of absences and foreign language use at home—added in Model 2 have very small negative effects on student performance controlling for the other variables. These are the number of absences the student had in the last month and whether people in the student's household frequently speak in another language.

Summary of Exploratory Study 3

The main finding of Exploratory Study 2 was that the faculty and staff that were interviewed perceived that the effects of student's social class and student's race were confounded.

The findings from Exploratory Study 3 indicate that race, class and gender (as currently measured) each have separate small, but statistically significant effects on student performance. These effects continue even when other potentially important measures are included in the model.

In addition to the effects race, gender and social class, the findings from Exploratory Study 3 indicate that students' perception of their grades, school transfers based on moves, and the number of books students' have access to in the home may affect student performance. The number of school absences and living in a foreign language speaking homes appears to have a small negative affect on student performance. The variables concerning length of residence were not significantly related to student performance.

SUGGESTIONS FOR FUTURE RESEARCH

Within the larger context of educational accountability, policy makers, educators and social scientists continue to look for new ways to increase student achievement and close any performance differences occurring between student groups. Accountability then requires some form of assessment to generate the needed information to see if these goals are being achieved. In Kentucky between 1991 and 1998, information gathering occurred by the KIRIS assessment.

The purpose of this research was to examine key parts of the KIRIS assessment to explore the extent to which there were performance differences by gender and race. The main findings of this study were that small academic performance differences by gender and race exist among Kentucky students. Specifically, women are somewhat more likely to have higher subject area scores than men in almost every subject area examined and at every grade level. These performance differences have decreased slightly and do not appear to be affected by geographic concentration. In addition, whites are more likely than blacks to have higher performance area scores. This gap has increased over the last six years. The concentration of African-Americans may affect this difference also in certain parts of the state. Other preliminary findings suggest that educators in a small sample of Kentucky's schools believed the racial performance differences were a result of students' social class; however, further quantitative analysis in the report calls that assumption into question.

At the end of Accountability Cycle 3, the state legislature abolished the KIRIS assessment. KDE is currently working to move local education agencies from KIRIS to a new system, the Commonwealth Accountability Testing System (CATS). This provides the state a new opportunity to refine its data collection procedures. Therefore, to better examine the effects of race, gender, class, test bias and other factors affecting students' academic performances, the following series of changes and studies is recommended.

Concepts, Variables, Reliability and Validity

It has become clear during this study that much of the data is needed for more rigorous examination of performance differences by race and gender are not currently available. Moreover, more valid and accurate measures are required before a more thorough and valid examination of performance differences by race, class and other factors can be done. The types of data necessary for the analysis can be collected several ways. The following are a series of research recommendations that stem specifically from conceptual, measurement, and analyses issues in generated during this study.

Throughout this study are several variables whose effects on student performance differences have yet to be appropriately examined because either no measure or a poor measure has been collected. Examples of several important variables that would be extremely useful for further research but need either more accurate and valid measurements are discussed below:

Individual Level Variables

Social Class - currently the only measure of social class and/or socioeconomic status collected at the student level is Title I participation. The problems with this measure as an indicator for social class and/or socioeconomic status were discussed in Exploratory Study 3. Including a variable indicating actual student participation in free and reduced price lunch would be a more accurate measure, if no measure of children's parent's socioeconomic status can be developed.

Mobility - The preliminary results of Exploratory Studies 2 and 3 suggest that understanding the effects of student mobility on performance is important.

Urbanity-Rurality - As noted in Exploratory Study 1, the concentration of individuals can have effects on students' academic performance. Previous research suggests that individuals in rural areas may perform differently on assessments; however, measures of ruralness or urbaness of individuals are only available aggregated to the school district level. Moreover, the federal department of education designated rurality not to individuals, but to schools based on their zip codes. A valid and reliable measure needs to be adopted.

Successful Transition to Adult Life - Currently the schools collect data on individual students' transition to adult life and aggregate the collected data to a school level measure; however, individual level information might be extremely useful in future performance difference studies.

Attendance - The results of Exploratory Study 3, and much of the earlier research, suggest that student attendance is positively related to academic performance.

School Level Variables

Curriculum - Curriculum is the core of education; however, the type of curricula and the effects of different curricula on student performance could not be examined in Exploratory Study 2.

Instruction - Because the effects of "tempo," "speed," and "language" on student performance, accurate measures need to be developed. The effects of these variables could not be examined in Exploratory Study 2.

Extended School Services and Family Resource Center programs - the preliminary findings of Exploratory Study 2 suggest that the programs offered may play a role in differential student performance and should be examined further. Currently there is no way to assess the effectiveness of different ESS and Family Resource Center models.

Dropout rates - the findings of this study and previous research suggest that there may be a differential dropout rate by race and gender. Therefore, dropout rates by race and gender need to be available for each school.

Retention rates - Like dropout rates, retention rates for schools are currently reported for inclusion in the non-cognitive measures of the academic index. This data should be reported also be reported by gender and race.

Proportion of African-American Faculty in School - Both Exploratory Study 1 and 2 suggest that racial concentration of students plays a role in student performance differences.

Parent Involvement - although school faculty often perceived that parent involvement was low in this study, other research suggests that this is an important variable to examine.

Leadership Style - KIRIS brought a new leadership style to school; its effects on student performance are yet unknown and could not be studied in Exploratory Study 2.

Number of Books in the Home - Number of Books in the Home is sometimes used as an indicator of individuals' ties to the larger society. In Exploratory Study 2, the findings were that varied experience, cultural enhancement, field trips, and enhanced technology were thought to be important factors related to high academic performance. Because this concept is so complex, a composite measure or a series of items needs to be developed and validated which may include these factors and others such as school's urban access and use of information sources such as newspaper, radio, TV, and technology.

County Level Variables

Dependency Indicators - the United States Department of Agriculture's (USDA) Economic Research Service has classified all the counties in the United States as to their dependency on certain sectors of the economy. These dependency indicators could be used as a filter to examine county level effects on student performance differences.

Shifts in Labor Market Sectors - Changes in a county's economic bases can create disruption within local communities and by that affect students. Measures of changes in labor market sectors within the community may be increasingly important as Kentucky moves from its current dependence on tobacco.

There are many other variables that would be helpful in the analysis of factors related to performance differences, but the collection of the individual level, school level, and county level variables on this list would make available data which could provide more valid analyses. There are two different ways this information can be collected. Specifically,

- 1) KDE can collect some of these individual student level measures by developing and including more reliable and valid measures on each student. Many

standardized and previously validated items are already available and could be easily included and/or substituted for items on the student questionnaire.

- 2) As a part of its KIRIS accountability formula, KDE collects a set of variables from local education agencies. These included concepts such as dropout rate, retention rate, rate of successful transition to adult life, and attendance rates. The procedure used by the state to collect these data could be easily modified to collect this information by race and gender. Starting in 1999, KDE began collecting dropout rates in this fashion for use in future research. The preliminary findings of Exploratory Study 2 would also suggest that mobility rate be collected as well.

Additional Research Initiatives

It is understandable that KDE does not currently have the resources or personnel that it would need to collect and analyze the information discussed in the previous section. Since the beginning of the KIRIS assessment, KDE has included two questionnaires as a part of the assessment process. Each student taking the KIRIS assessment fills out a questionnaire; their teachers also have another on which to respond. While some of these items on these questionnaires would be useful in examining the relationships between performance and other factors, most of this data has not been analyzed.

In addition, free or inexpensive data sets collected by other state and federal institutions are also available and can easily be connected to the KIRIS and/or CATS data. Two organizations that have datasets that might be useful are the National Center for Education Statistics (NCES) and the USDA. The NCES School District Data Book contains information from the decennial census aggregated to the school district level and the USDA's 1997 Census of Agriculture has contextual information about farm and rankings of farm production.

Qualitative and Quantitative Research

To assess the causes of and solutions for the performance differences among students, a series of qualitative and quantitative research studies that can be cross-validated is fundamental. The current data collected on student and teacher questionnaires, from local education agencies and from other sources can then also be used for longitudinal and cross-sectional research.

Qualitative

Qualitative research attempts to understand the causes of social behavior by comparing and contrasting the detailed views of informants along with first-hand observations of behavior (Creswell 1994). As such, qualitative research is very useful for generating theories of behavior. Currently, many hypotheses relate a variable or variables to performance differences; however, there are no general conceptual schemes available to integrate these hypotheses into a theory of why performance differences occur. It will take valid and reliable qualitative research to generate this theory.

The usefulness of qualitative research relies on the ability of the researcher to generalize from the specific cases that he or she examines to a larger population. To improve the validity and generalizability of qualitative studies, it is necessary to draw a sample of schools that “provides simultaneous maximization or minimization of both the differences and the similarities of data” (Glaser and Strauss 1967:55). Thus to generate a good theoretical understanding about the causes of any performance differences found in Kentucky’s assessment data, in a qualitative sample maximizing the differences in the schools sampled will be necessary. Specifically, there are schools with least one of four different performance difference trends, which should be included in the sample to improve validity and generalizability. A few schools with each of the following performance trends need to be included in any qualitative study:

- 1) Schools with consistently small performance differences using the variables of interest.
- 2) Schools with performance differences that diminish over a time using the variables of interest.
- 3) Schools with performance differences that were consistently large using the variables of interest.
- 4) Schools with performance differences have increased over time using the variables of interest.

The schools sampled will need to have sufficient numbers of individuals in all population subgroups under investigation to make sure the estimates of performance differences are stable. As an additional attempt to increase generalizability, the selection procedure should also ensure geographic and grade level diversity.

Quantitative

While qualitative research is very useful for generating new theories concerning the causes of behavior, quantitative research is concerned with taking a theory and developing numerical indicators of concepts in that theory such that statistical testing can be performed. The section above enumerates several concepts that can be measured and examined using quantitative research. The use of statewide questionnaires, such as the KIRIS student question and the KIRIS teachers’ questionnaire, or interviews can give researchers many variables to test hypotheses about student performance differences if valid measures are used.

Why Triangulate Qualitative and Quantitative Research?

While quantitative and qualitative research are often performed separately, when both of these research designs are used to explore the same questions, one is able to test hypotheses generated by the qualitative research using quantitative measures and analysis. This process is called cross-validation or triangulation (Denzin 1989). Cross-validation strengthens the reliability and validity of both quantitative and qualitative analysis. In order to cross-validate research on student performance differences, the qualitative should be performed before the survey construction begins, unless it is possible to begin both processes with a clear idea of the research questions are to be examined.

The Need for Organization

The passage of KERA in 1990 put Kentucky at the forefront of a new wave of education reform. The Commonwealth's prominent role in the education reform movement has led to a plethora of research on each component of the reform by KDE, other educators, social scientists, and various statewide and national associations and foundations. There is also a body of research studies that compares aspects of Kentucky's reform initiatives and those of other states. However, no location houses the entire canon of research on Kentucky's educational reform. Currently individuals wishing to do research on Kentucky's education reform may only have access to resources available at the libraries in their region or through personal contacts with other researchers. The same is true of ongoing research initiatives. This has implications both for the comparability of research and for the resource use. Some type of clearinghouse would help to address these issues. Educators and policy makers also need information with which to decide the implementation of the education reform. It is therefore important that they can access existing information quickly and easily. The characteristics of such a clearinghouse or consortium would need to be determined, however, any organization of the existing data and theory on the KIRIS assessment would facilitate future research.

REFERENCES

- AAUW (American Association of University Women). 1992. How Schools Shortchange Girls: A Study of Major Findings of Girls and Education. Washington, DC: American Association of University Women.
- Alexander, Karl L., Doris R. Entwisle, and Susan L. Dauber. 1996. "Children in Motion: School Transfers and Elementary School Performance." Journal of Educational Research 90(1):3-12.
- Alfaro, Robert. 1999. "The Technology-Reading Connection." Educational Leadership 56(6):48-50
- Banks, Cherry A. McGee and James A. Banks. 1997. "Reforming Schools in a Democratic Pluralistic Society." Educational Policy 11(2):183-193.
- Banks W. Curtis, Gregory V. McQuarter, and Janet L. Sonne. 1995. A Deconstructive Look at the Myth of Race and Motivation. Journal of Negro Education. 64(3):307-325.
- Baron, J.N. and J. Pfeffer. 1994. "The Social Psychology of Organizations and Inequality." Social Psychology Quarterly 57(3):190-209.
- Bjorgvinsson, Throstur and Paula Kerr. 1995. "Use of a Common Language Effect Size Statistic." American Journal of Psychiatry 152(1):151.
- Boone, William J. et.al. 1997. "Race, Gender, Test Length, and Missing Data: Why Estimates of Performance May Be Clouded." Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (March 22, 1997).
- Bureau of Economic Analysis. 1997. "Bearfacts: Fayette, Kentucky 1987-97." <http://cbpa.louisville.edu/ksdc/sdc/bearfacts/bf067.txt> Accessed on July 15, 1999.
- Calabrese, Raymond L. and John Poe. 1991. "Alienation: An Explanation of High Dropout Rates among African American and Latino Students." Educational Research Quarterly 14(4):22-26.
- Campbell, Donald T. 1984. "Can We be Scientific in Applied Social Science?" Chapter 1 in Evaluation Studies Review Annual. 9:26-48.
- Catsambis, S. 1994. "The Path to Math: Gender and Racial-Ethnic Differences in Mathematics Participation from Middle School to High School." Sociology of Education 67: 199-215.
- , 1995. "Gender, Race, Ethnicity, and Science Education in the Middle Grades." Journal of Research in Science Teaching 32(3):243-257.

- Chandler, Louis. 1997. Forced Busing: A Staff Report of the Allegheny Institute for Public Policy. Pittsburgh, PA: Allegheny Institute.
- Clark, M.L. 1991. "Social Identity, Peer Relations, and Academic Competence of African-American Adolescents." Education and Urban Society 24(1):41-52.
- Clewell, B.C. and B. Anderson. 1991. Women of Color in Mathematics, Science, and Engineering: A Review of the Literature. Washington, DC: Center for Women Policy Studies.
- Cohen, Jacob. 1977. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press.
- Cohen, L.L. and J.K. Swim. 1995. "The Differential Impact of Gender Ratios on Women and Men: Tokenism, Self-Confidence, and Expectations." Personality and Social Psychology Bulletin 21(9):876-884.
- Collins-Eaglin, Jan and Stuart A. Karabenick. 1993. "Devaluing of Academic Success by African-American Students: On "Acting White" and "Selling Out." Paper presented at the annual Meeting of the American Educational Research Association. Atlanta, GA (April 12-16, 1993).
- Cook, Peggy J., and Mizer, Karen L. 1989. "The revised ERS County Typology: An Overview." Rural Development Report 89. Washington, DC: USDA Economic Research Service, Rural Economy Division.
- Cotton, Kathleen. 1996. Affective and Social Benefits of Small-Scale Schooling. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools.
- Creswell, John W. 1994. Research Design: Qualitative and Quantitative Approaches. Thousand Oaks, CA: Sage.
- Cross, Ray, et. al. 1996. "Do the Effects of Effective Elementary Schools Endure?" Paper presented at the Annual Meetings of the American Educational Research Association. New Orleans, LA (April 4-8, 1994).
- Darling-Hammond, L. 1992. Standards of Practice for Learner-Centered Schools. New York: National Center for Restructuring Education, Schools and Teaching, Columbia University.
- Denzin, Norman K. 1989. The Research Act: A Theoretical Introduction to Sociological Methods. Englewood Cliffs, NJ: Prentice-Hall.
- Education Trust. 1999. Dispelling the Myth: High Poverty Schools Exceeding Expectations. Washington, DC: Education Trust.

- Entwisle, D.R. and K.L. Alexander. 1992. "Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School." American Sociological Review 57(1):72-84.
- Evans, M.O. 1992. "An Estimate of Race and Gender Role-Model Effects in Teaching High School." Journal of Economic Education 23(3):209-217.
- Fernanda Fontinas, Ana M. Morais, and Isabel P. Neves. 1995. "Students' Coding Orientation and School socializing Context in Their Relation with Students' Scientific Achievement." Journal of Research in Science Teaching 32:445-62..
- Finn, J.D. and D.A. Rock. 1997. "Academic Success Among Students at Risk for School Failure." Journal of Applied Psychology 82(2):221-234.
- Foley, Douglas E. 1991. "Reconsidering Anthropological Explanations of Ethnic School Failure." Anthropology and Education Quarterly 22(1):60-86.
- Fordham, S. 1988. "Racelessness as a Factor in Black Students' School Success: Pragmatic Strategy or Pyrrhic Victory?" Harvard Education Review 58(1):54-84.
- Friedman, Herbert. 1968. "Magnitude of Experimental Effect and a Table for Its Rapid Estimation." Psychological Bulletin 70(4):245-251.
- Frisby, Craig L. 1998. "Poverty and Socioeconomic Status." Pp. 241-270 in Test Interpretation and Diversity: Achieving Equity in Assessment, edited by J. Sandoval, C.L. Frisby, K.F. Geisinger, J.D. Scheuneman, and J.R. Grenier. Washington, DC: American Psychological Association.
- Glaser, Barney G. and Anselm L. Strauss. 1967. The Discovery of Grounded Theory: Strategies for Qualitative Research. New York: Aldine de Gruyter.
- Gollnick, Donna M. and Philip C. Chinn. 1994. Multicultural Education in a Pluralistic Society Columbus, OH: Merrill.
- Greenwood, Charles R., Barbara Terry, and Cheryl A. Utey. 1993. "Achievement, Placement, and Services: Middle School Benefits of Classwide Peer Tutoring Used at the Elementary School." School Psychology Review 22(3):497-516.
- Haase, Richard F., Michael V. Ellis, and Nicholas Ladany. 1989. "Multiple Criteria for Evaluating the Magnitude of Experimental Effects." Journal of Counseling Psychology 36(4):511-516.
- Haney, W. and G. Madaus. 1989. "Searching for Alternatives to Standardized Tests: Whys, Whats, and Whithers." Phi Delta Kappan 70(9):683-687.

- Hays, Samuel P. 1989. Beauty, Health and Permanence: Environmental Politics in the United States, 1955-1985. Cambridge University Press.
- Herman, Joan, Pamela Acshbacher, and Lynn Winters. 1992. A Practical Guide to Alternative Assessment. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hooijberg, R. and N. DiTomaso. 1996. "Leadership In and Of Demographically Diverse Organizations." Leadership Quarterly 7(1):1-19.
- Howell, F.M., Y. Tung, and C. Wade-Harper. 1996. The Social Cost of Growing-Up in Rural America: Rural Development and Social Change during the Twentieth Century. Bulletin of the Mississippi Agricultural and Forestry Experiment Station.
- Howley, Craig. 1996. Ongoing Dilemmas of School Size: A Short Story. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools.
- Jackson, P.B., P.A. Thoits, and H.F. Taylor. 1995. "Composition of the Workplace and Psychological Well-Being: The Effects of Tokenism on America's Black Elite." Social Forces 74(2):543-557.
- Jencks, Christopher and Meredith Phillips. (Eds.) 1998. The Black-White Test Score Gap. Washington, D.C.: The Brookings Institution Press.
- Kailin, J. 1994. "Anti-Racist Staff Development for Teachers: Considerations of Race, Class, and Gender." Teaching and Teacher Education 10(2):169-184.
- Kanter, R.M. 1977. Men and Women of the Corporation. New York: Basic Books.
- KDE (Kentucky Department of Education). 1995. KIRIS Accountability Cycle 1 Technical Manual. Frankfort, KY: Kentucky Department of Education.
- , 1997. KIRIS Accountability Cycle 2 Technical Manual. Frankfort, KY: Kentucky Department of Education.
- KDE (Kentucky Department of Education), Division of Curriculum Development, Office of Academic and Professional Development. 1999a. "Core Content for Assessment." <http://www.kde.state.ky.us/oapd/curric/Publications/CoreContent/>. Accessed on July 18, 1999.
- KDE (Kentucky Department of Education), Division of Program Resources, Office of Special Instruction Services. 1999b. "Title I, Part A Program Requirements." http://www.kde.state.ky.us/osis/resources/t1/title_1_requirements.asp. Accessed on July 8, 1999.

- KDE (Kentucky Department of Education), Division of School Improvement. 1999c. "Kentucky Leadership Academy." http://www.kde.state.ky.us/olsi/improve/kla/kla_program_focus.asp. Accessed on July 18, 1999.
- Keeves, J.P. and L.J. Saha. 1994. "Measurement of Social Background." Pp. 3715-22 in The International Encyclopedia of Education, Vol. 7, 2nd ed. New York: Elsevier Science.
- Keith, N.Z. 1996. "A Critical Perspective on Teacher Participation in Urban Schools." Educational Administration Quarterly 32(1):45-79.
- Kelley, Carolyn. 1998. "The Kentucky School-Based Performance Award Program: School Level Effects." Educational Policy 12(3):305
- Kennedy, E. 1995. "Correlates of Perceived Popularity Among Peers: A Study of Race and Gender Differences Among Middle School Students." Journal of Negro Education 64(2):186-195.
- Kleinfeld, Judith. 1998. The Myth that Schools Shortchange Girls: Social Science in the Service of Deception. Washington, DC: Women's Freedom Network.
- Kohr, R.L., J.R. Masters, J.R. Coldiron, R.S. Blust, and E.W. Skiffington. 1989. "The Relationship of Race, Class, and Gender with Mathematics Achievement for Fifth-, Eighth-, and Eleventh-Grade Students in Pennsylvania Schools." Peabody Journal of Education 66(2):147.
- Kulm, G. and S. M. Malcolm. Science Assessments in the Service of Reform. Washington, DC: American Association for the Advancement of Science.
- Lawrence, B.S. 1997. "The Black Box of Organizational Demography." Organization Science 8(1):1-22.
- Makedon, Alexander. 1992. "Is Alice's World Too Middle Class? Recommendations for Effective Schools Research." <http://ericae.net/ericdb/ED346612.htm> Accessed on July 17, 1999.
- Mathews, W. 1984. "Influences on the Learning and Participation of Minorities in Mathematics." Journal of Research in Mathematics Education 15:84-95.
- McDowell, Ceasar L. 1990. "The Unseen World: Race Class and Gender Analysis of Science Education Research." Journal of Negro Education 59(3):273-291.
- McGraw, Kenneth O. and S.P. Wong. 1992. "A Common Language Effect Size Statistic." Psychological Bulletin 111(2):361-5.

- MSBE (Michigan State Board of Education). "Michigan State Board of Education: Professional Development Definition and Standards."
http://www.mde.state.mi.us/off/s_prg/pdstandards.shtml Accessed on July 18, 1999.
- Murray, Carolyn B. 1996. "Estimating Achievement Performance: A confirmation Bias." Journal of Black Psychology. 22(1)67-85.
- NCES (National Center for Educational Statistics). 1994. School District Data Book CD-Rom. Washington, DC: U.S. Department of Education. Office of Educational Research and Improvement.
- , 1998. Dropout Rates in the United States: 1996.
<http://nces.ed.gov/pubs98/dropout/> Accessed on July 5, 1999.
- Nweze, Billie R. 1993. Increasing Parent Involvement, Student Attendance and Appropriate School Behavior of At-Risk Middle School Students through Parent Partnerships.
<Http://ericae.net/ericdb/ED366485.htm> Accessed on July 15, 1999.
- Olson, L. 1997. "Quality Counts: Examining Race and Demography."
<http://www.edweek.org/qc/intros/race.htm> Accessed December 1997.
- Orenstien, A.C. and D.U. Levine. 1990. "Social Class, Race, and School Achievement: Problems and Prospects." Journal of Teacher Education 40(5):17.
- Orfield, Gary and John T. Yun. 1999. "Resegregation in America's Schools." A Report of the Civil Rights Project. Harvard University.
<http://www.law.harvard.edu/groups/civilrights/publications/resegregation99.html>
 Accessed on July 9, 1999.
- Page, S. and R. Rosenthal. 1990. "Sex and Expectations of Teachers and Sex and Race of Students as Determinants of Teaching Behavior and Student Performance." The Journal of School Psychology 28:119-131.
- Parker, S., S. Greer, and B. Zuckerman. 1988. "Double Jeopardy: The impact of poverty on Early Child Development." Pediatric Clinics of North America 35:1227-40.
- Plecki, M. 1991. "The Relationship Between Elementary School Size and Student Achievement." Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. (ERIC/CRESS Accession No. RC 019 440).
- Porter, Maureen K. 1996. "Moving Mountains: Reforms, Resistance and Resiliency in an Appalachian Kentucky High School. Rural Educator 18:25-9.

- Quellmalz, Edys, Patrick M. Shields and Michael S. Knapp. 1995. School Based Reform - Lessons from a National Study: A Guide for School Reform Teams. Washington, D.C. U.S. Office of Education.
- Raywid, M. 1996. Taking Stock: The Movement to Create Mini-Schools, Schools-Within-Schools, and Other Small Schools. New York: ERIC Clearinghouse on Urban Education.
- Retherford, R.D. 1975. The Changing Differentials in Mortality. Westport: Greenwood Press.
- Revis, Emma. 1995. Schools Achieving Gender Equity (ERIC Accession No. ED382785).
- Roscigno, Vincent J. 1999. "The Black-White Achievement Gap, Family-School Links and the Importance of Place. Sociological Inquiry 69(2)159-86.
- Rosenthal, R. 1973. On the Social Psychology of the Self-Fulfilling Prophecy: Further Evidence for Pygmalion Effects and their Mediating Mechanisms. New York: MSS Modular Publications, Module 53, 1-28.
- Rowley, Stephanie J. 1996. "Perceptions of Discrimination and Achievement in African American Students. Paper presented at the annual Meeting of the American Psychological Association (Toronto, Ontario, Canada, August 9-13 1995).
- Samejima, F. 1969. "Estimation of Latent Ability Using a Response Pattern of Graded Scores." Psychometrika Monograph Supplement, No. 17.
- Schack, Gina D., Ellen McIntyre, and Diane Wells Kyle. 1993. "Equity In Kentucky's Primary Programs." Equity and Excellence in Education. 26(3):37-41.
- Sechrest, Lee and William H. Yeaton. 1982. "Magnitudes of Experimental Effects in Social Science Research." Evaluation Review 6(5):579-600.
- Sleeter, Christine E. 1992. "Resisting Racial Awareness: How Teachers Understand the Social Order from their Racial Gender and Social Class Locations." Educational Foundations 6(2): 7-32.
- SSTA (Saskatchewan School Trustees Association) 1999. "Curriculum." <http://www.ssta.sk.ca/research/curriculum/curriculum.htm>
Accessed on July 17, 1999.
- Stanfield, John H, II and Dennis M. Rutledge. (Eds.) 1993. Race and Ethnicity in Research Methods. Newbury Park, CA: Sage.
- Steele, C.M. 1992. "Race and the Schooling of Black Americans." The Atlantic 269(4):68.

- Swanson, L.L. (editor). 1996. Racial/Ethnic Minorities in Rural Areas: Progress and Stagnation, 1980-1990. Rural Economy Division, Economic Research Service, U.S. Department of Agriculture. Agricultural Economic Report No. 731.
- Tatum, B.D. 1992. "Talking about Race, Learning about Racism: The Application of Racial-Identity Development Theory in the Classroom." Harvard Educational Review 62(1):1-24.
- Trusty, Jerry and Katherine Dickey. 1993. "Alienation from School: An Exploratory Analysis of Elementary and Middle School Students." Journal of Research and Development in Education 26(4):232-242.
- Tucker, Carolyn M., Yvette R. Harris, and Beverly A. Brady. 1996. "The Association of Selected Parent Behaviors with the Academic Achievement of African -American and European American Children." Child Study Journal 26(4)253-277.
- U.S. Department of Education. 1996. "Targeting, Formula, and Resource Allocation Issues: Focusing Federal Support Where the Needs Are Greatest." <http://www.ed.gov/offices/OUS/eval/esed/targetan.html>
Accessed on July 10, 1999.
- U.S. Census Bureau. 1999. "Metropolitan Areas and Components, 1999, with FIPS Codes." <http://www.census.gov/population/estimates/metro-city/99mfips.txt>. Accessed July 14, 1999.
- Valli, L. 1995. "The Dilemma of Race: Learning to be Color Blind and Color Conscious." Journal of Teacher Education 46(2):120-129.
- Weis, Lois and Michelle Fine (Eds.) 1993. Beyond Silenced Voices: Class, Race and Gender In United States Schools. Albany, NY: State University of New York Press.
- Wetter, Robert E. 1999. Personal Communication. July 19, 1999.
- Wilkinson, Kenneth P. 1991. The Community in Rural America. New York: Greenwood Press.
- Wilson, William Julius. 1987. The Truly Disadvantaged: The Inner City, The Underclass, and Public Policy. Chicago: University of Chicago Press.
- Wright, Doris J. 1993. "Multicultural Issues in Organizational Consultation: A Conceptual Model for Intervention." Paper presented at the Annual Meeting of the American Psychological Association. (Toronto, Ontario, Canada, August 20-24, 1993)
- Yoder, J.D. 1994. "Looking Beyond Numbers: The Effects of Gender Status, Job Prestige, and Occupational Gender-Typing on Tokenism Processes." Social Psychology Quarterly 57(2):150-159.

Zill, N., K.A. Moore, E.W. Smith, T. Stief, and M.J. Coiro. 1991. The Life Circumstances and Development of Children in Welfare Families: A Profile Based on National Survey Data. Washington, DC: Child Trends.

Appendix A⁴¹

KIRIS Cycle 2 and Cycle 3 Subject Area Performance as
Measured by the Theta scale for All Grades Tested by
Gender

⁴¹ Data Tables were added to all Figures in Appendix A by KDE.

Figure A1 KIRIS Reading Theta for Grade 4 by Gender, 1993-1998

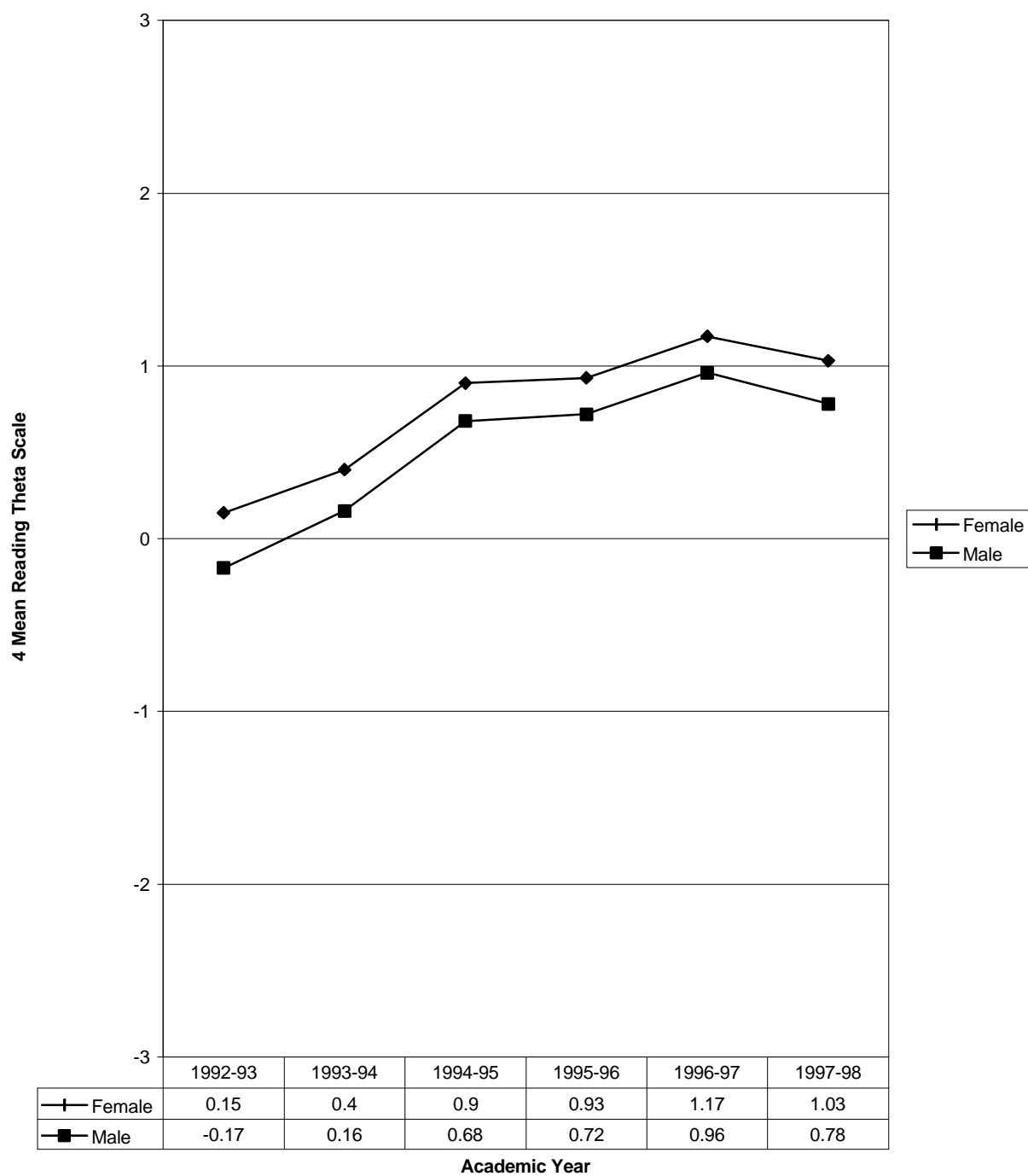


Figure A2. KIRIS Mathematics Theta for Grade 4/5 by Gender, 1993-1998

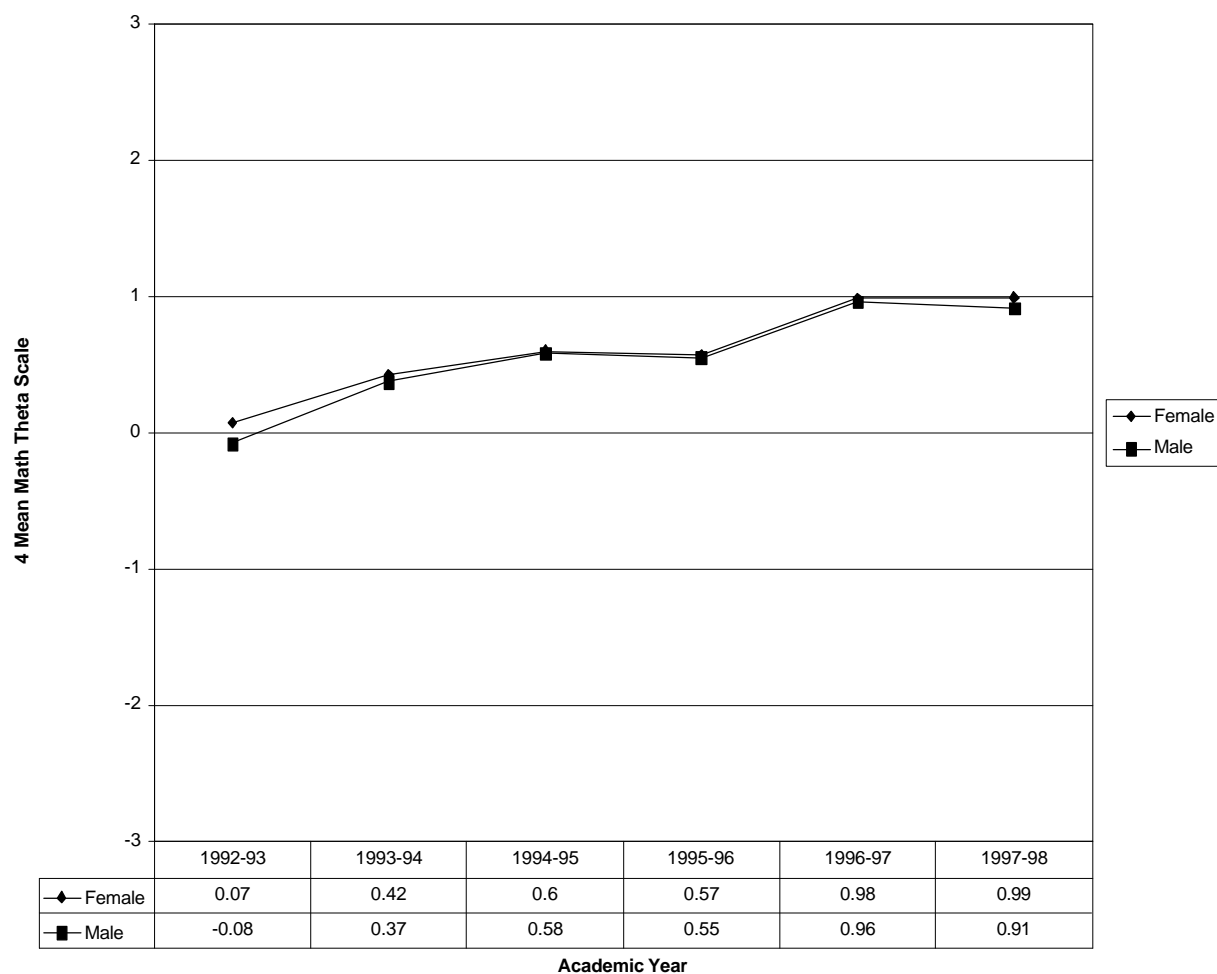


Figure A3. KIRIS Science Theta for Grade 4 by Gender, 1993-1998

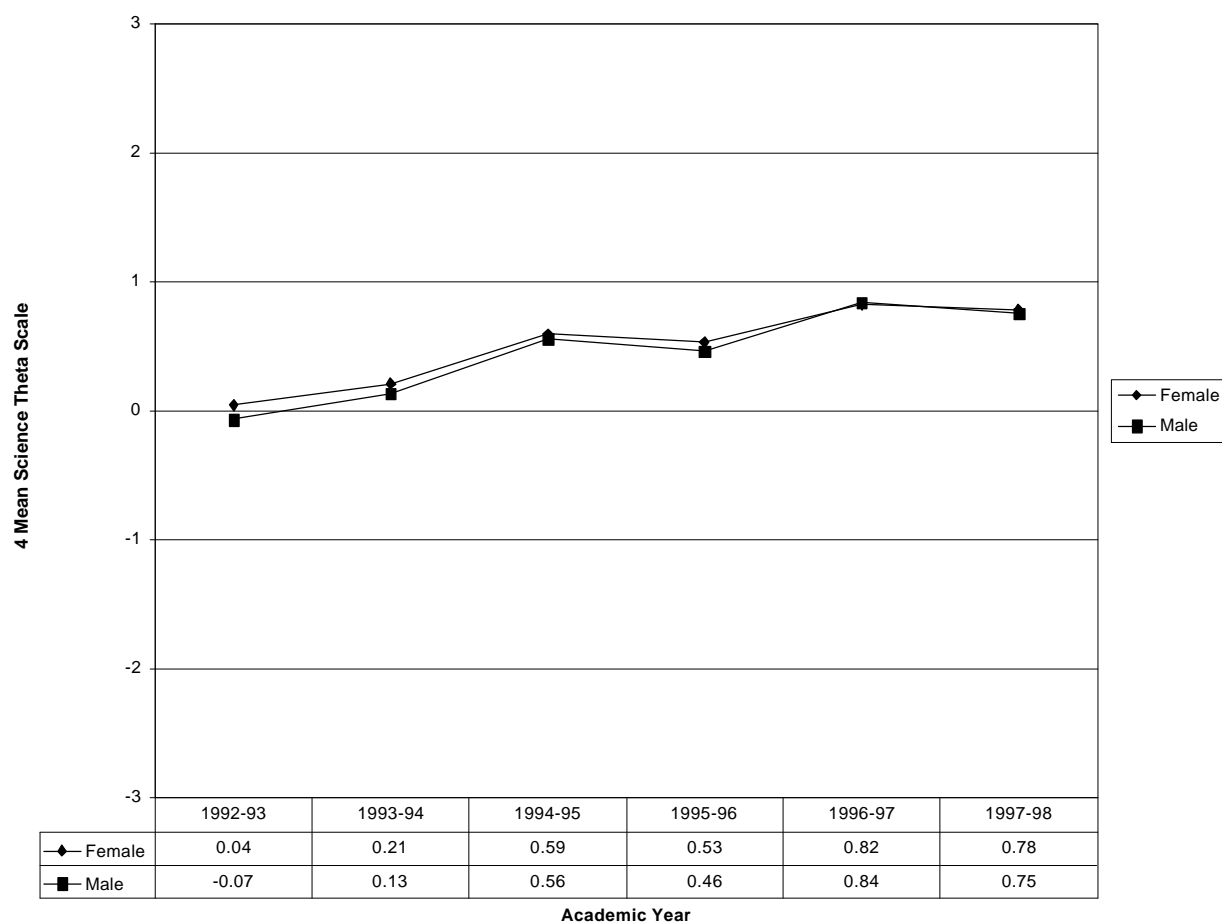


Figure A4. IRIS Social Studies Theta for Grade 4/5 by Gender, 1993-1998

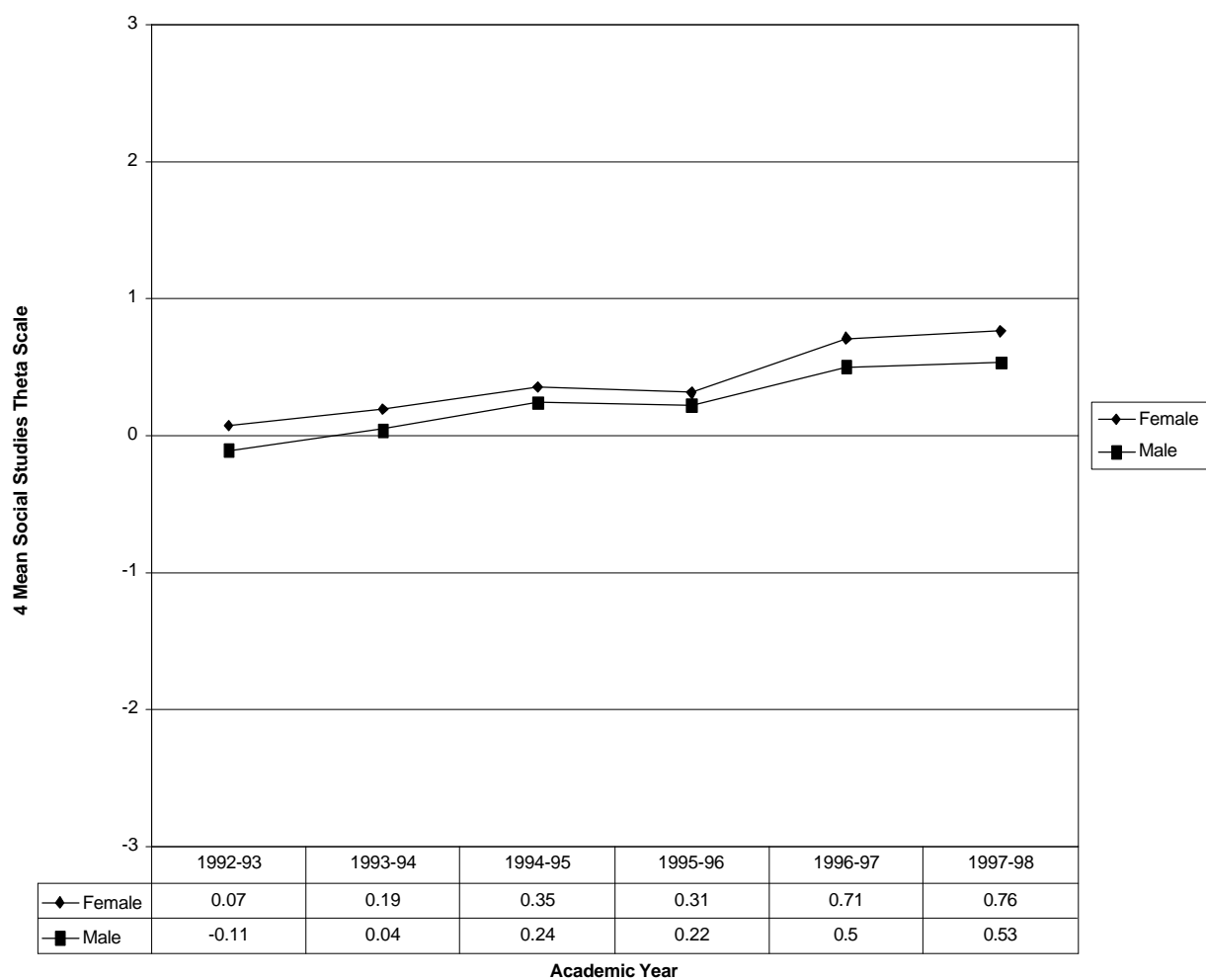


Figure A5. KIRIS Reading Theta for Grade 7/8 by Gender, 1993-1998

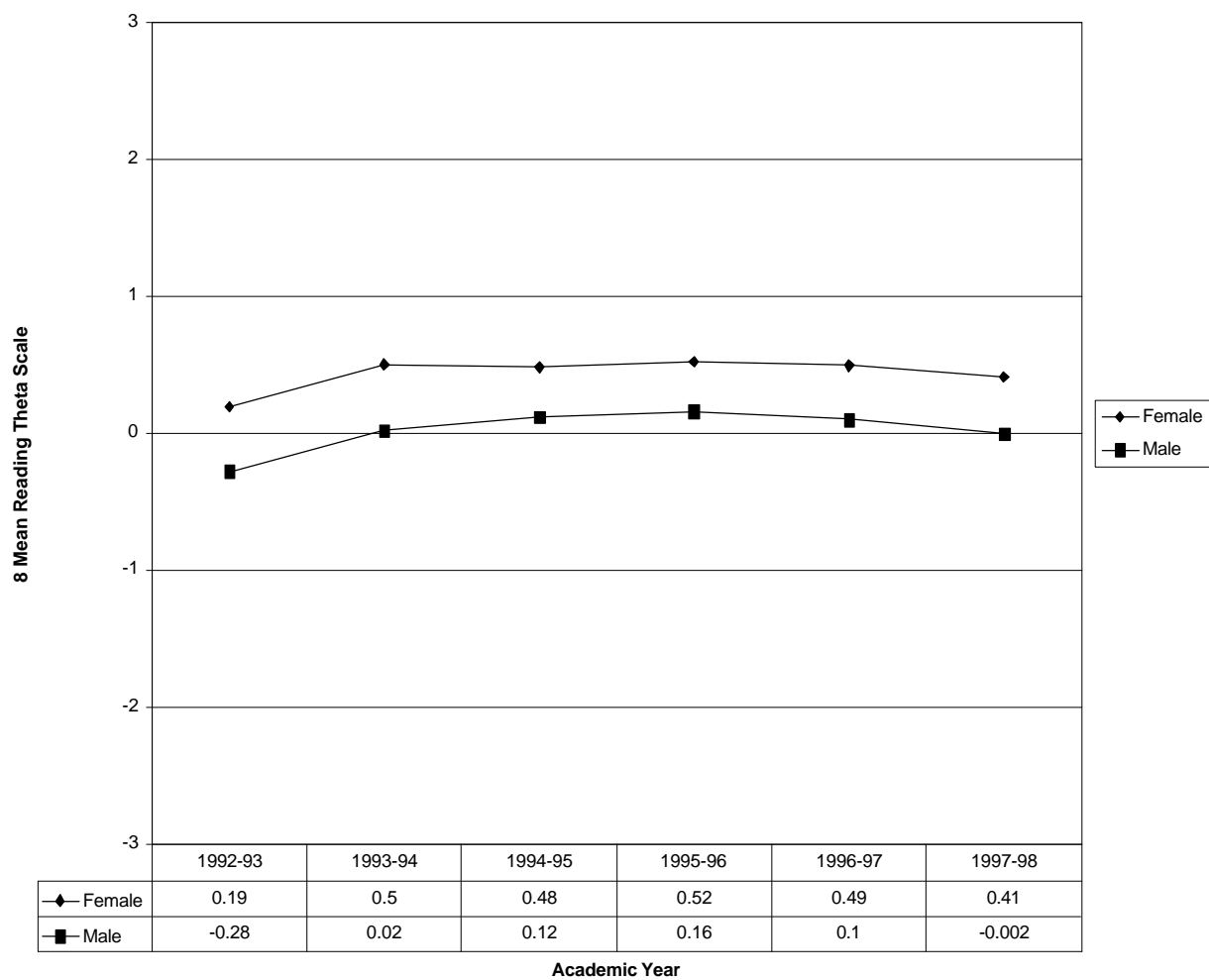


Figure A6. KIRIS Mathematics Theta for Grade 8 by Gender, 1993-1998

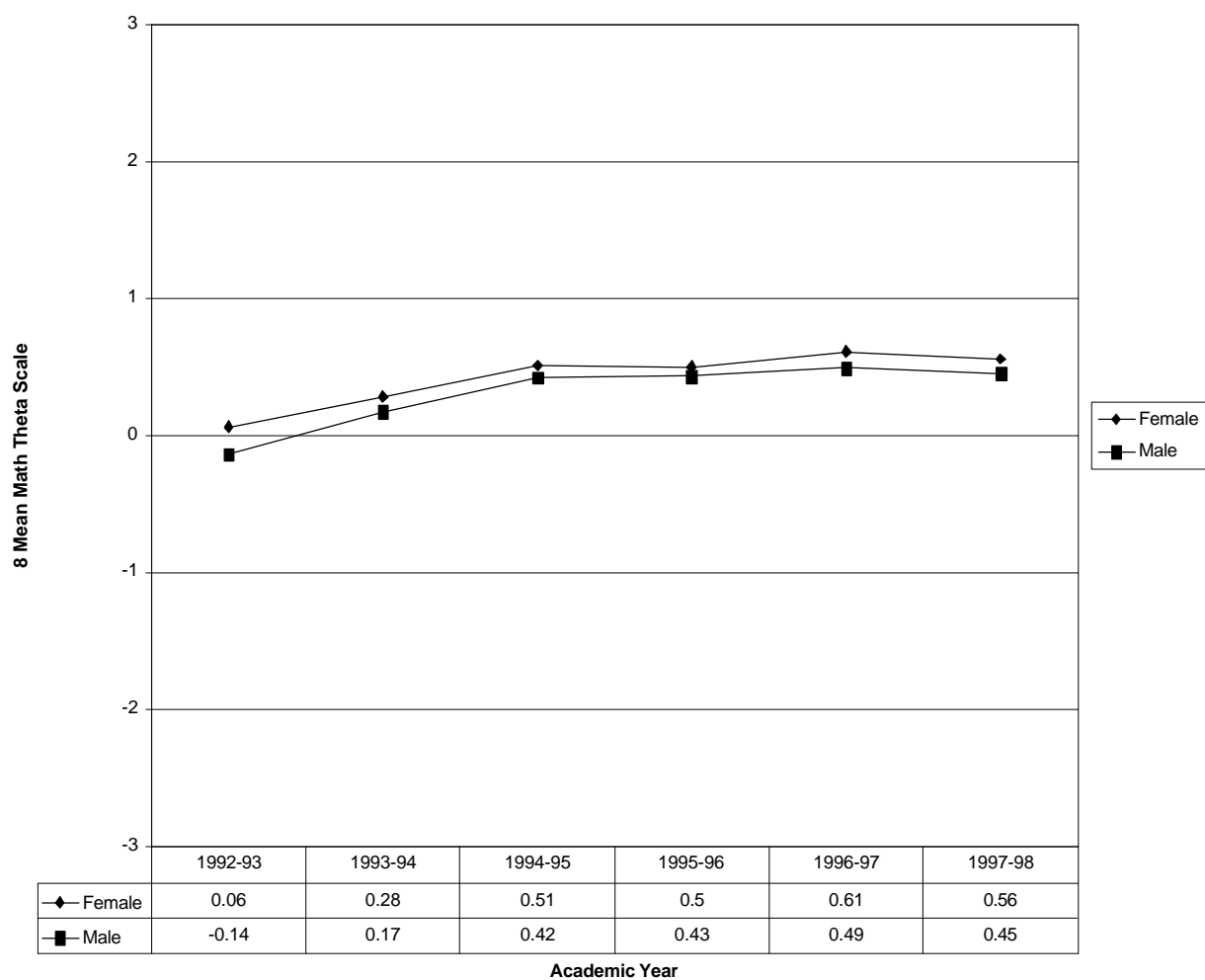


Figure A7. KIRIS Science Theta for Grade 7/8 by Gender, 1993-1998

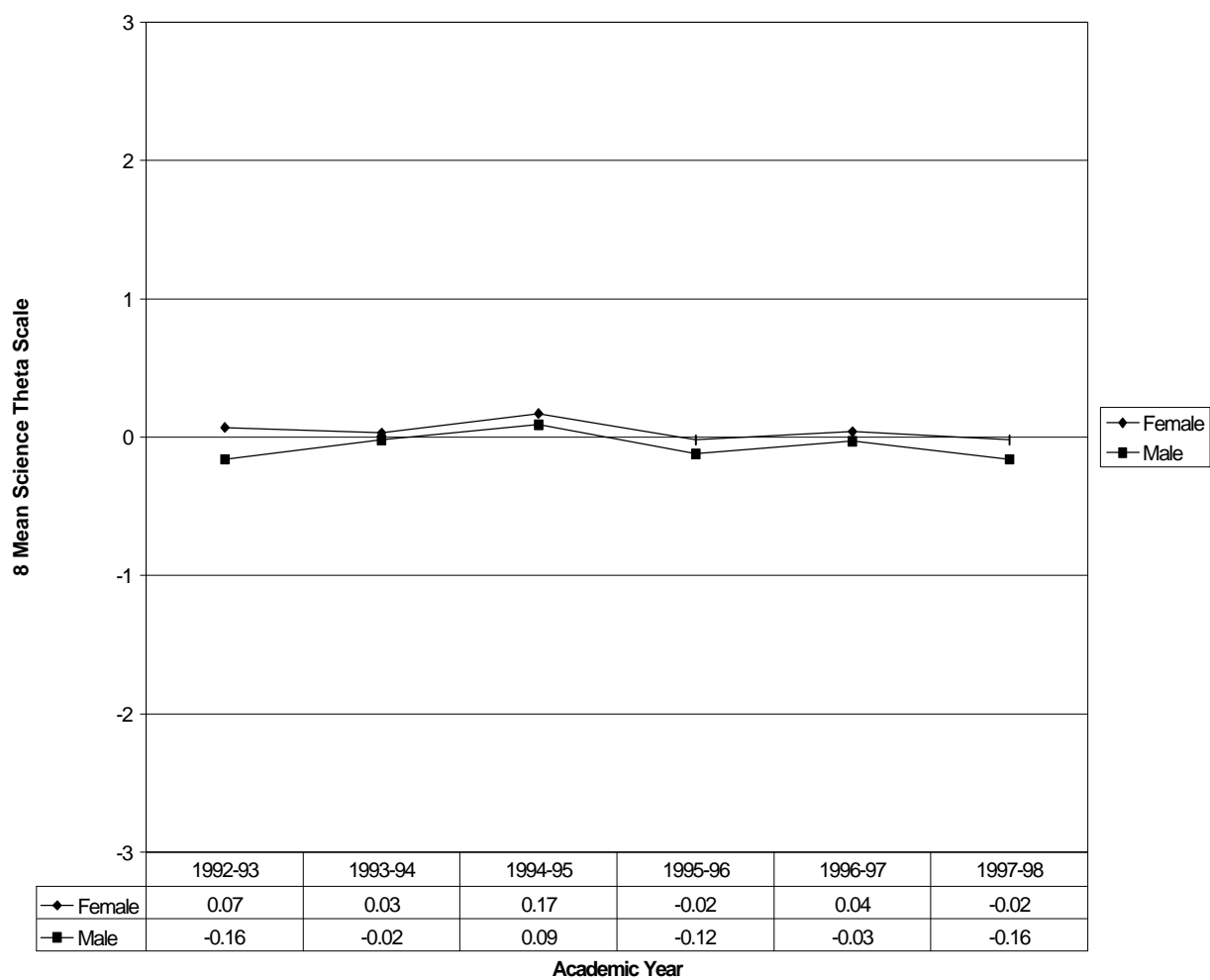


Figure A8. KIRIS Social Studies Theta for Grade 8 by Gender, 1993-1998

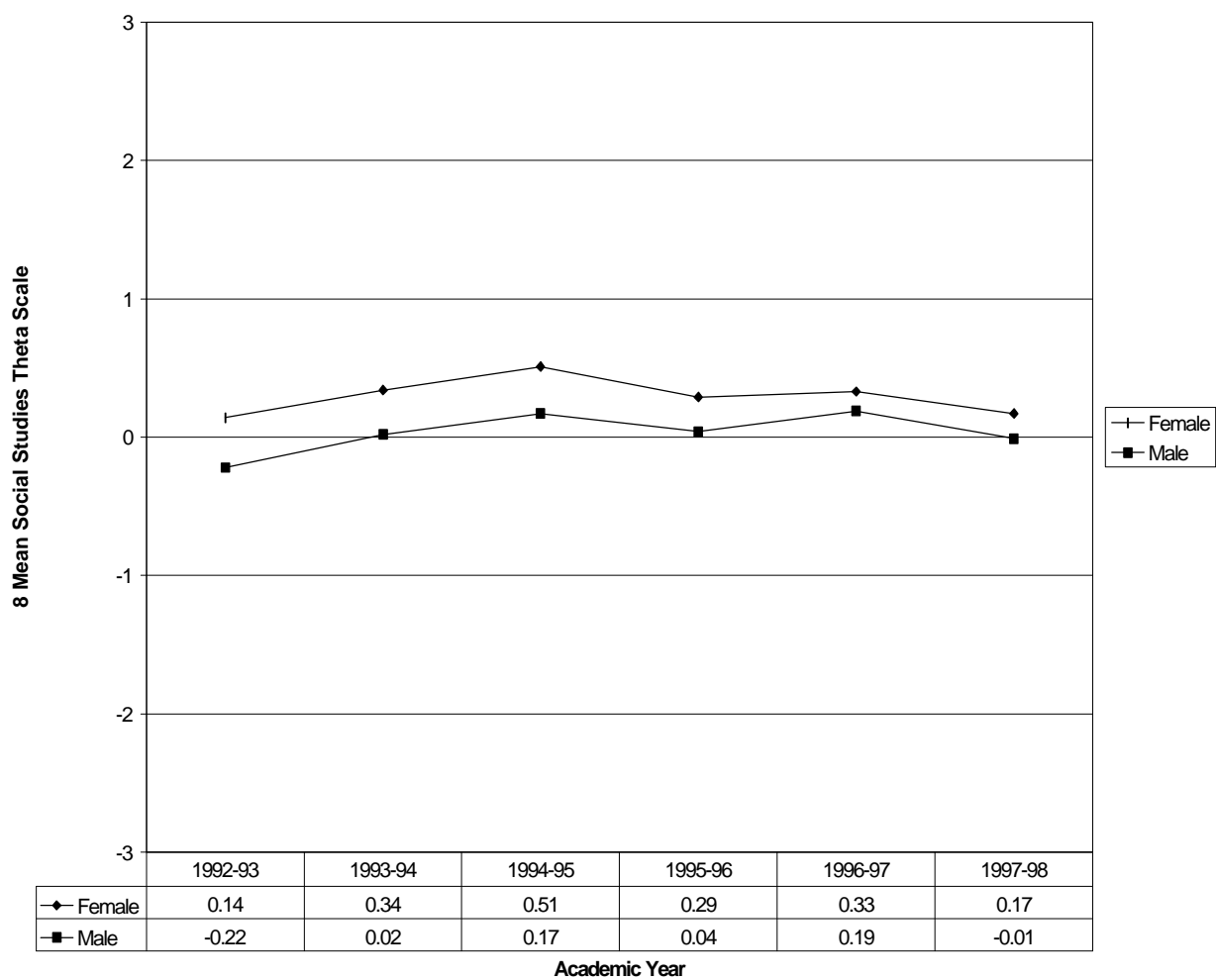


Figure A9. KIRIS Reading Theta for Grade 11/12 by Gender, 1993-1998

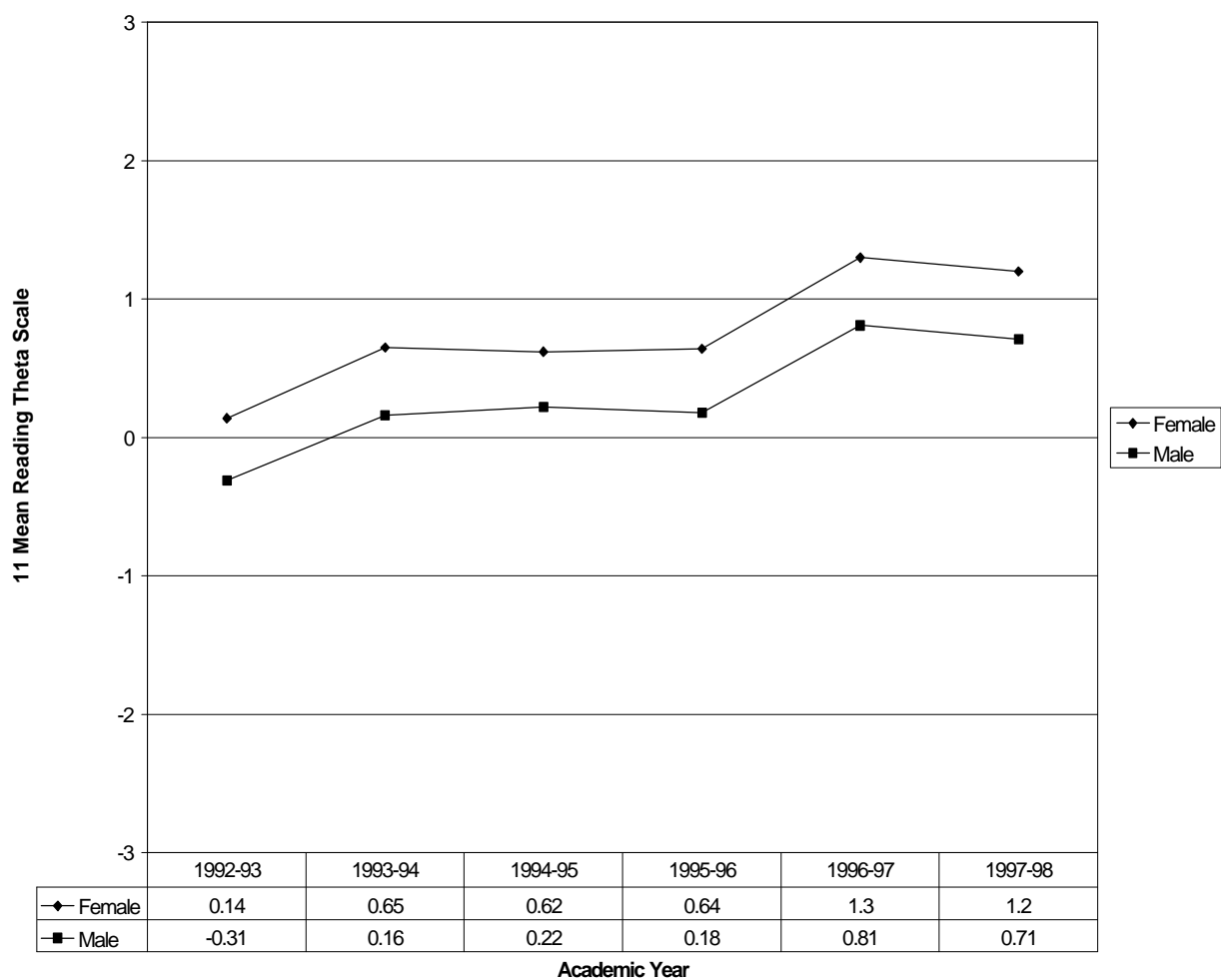


Figure A10. KIRIS Mathematics Theta for Grade 11/12 by Gender, 1993-1998

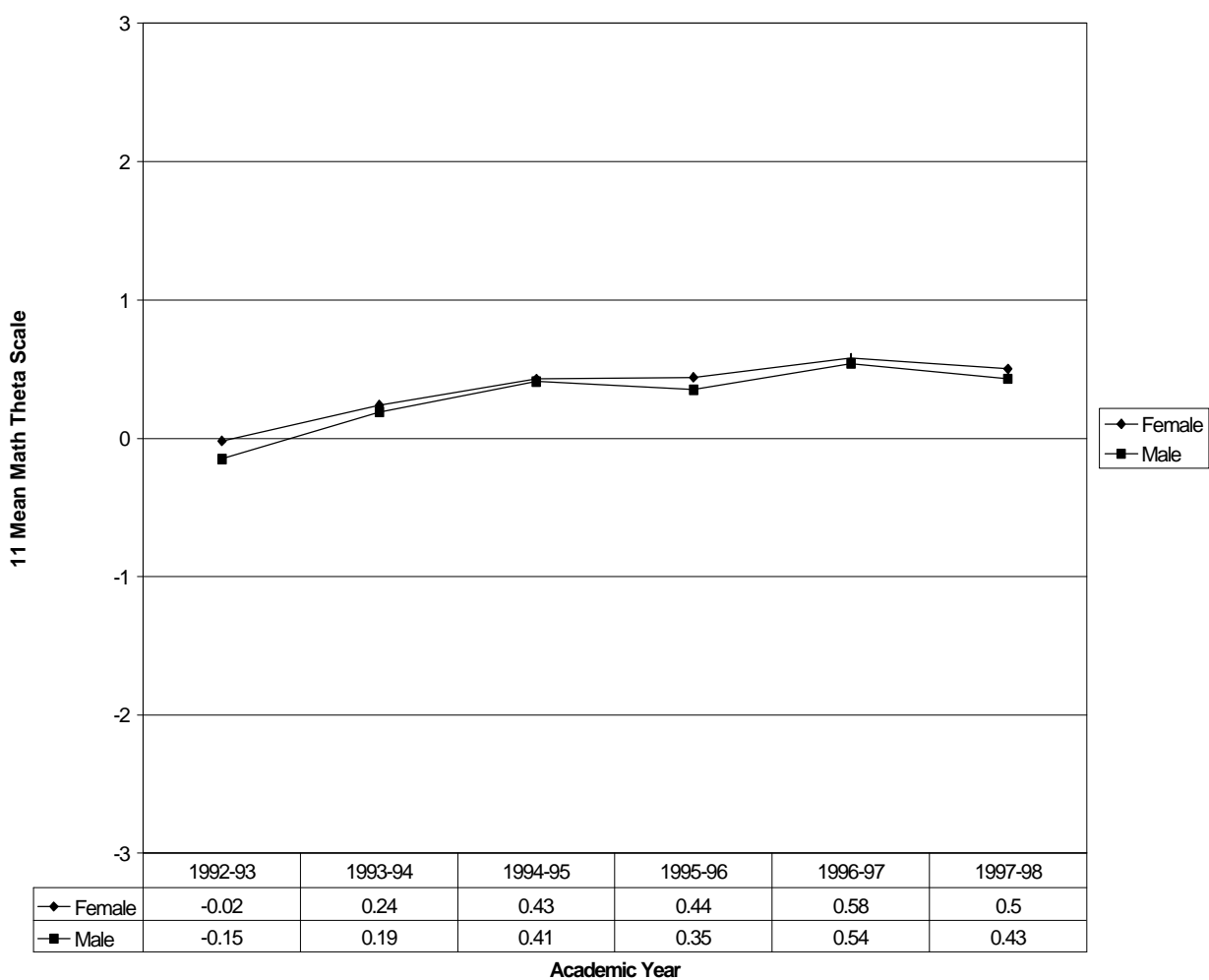


Figure A11. KIRIS Science Theta for Grade 11/12 by Gender, 1993-1998

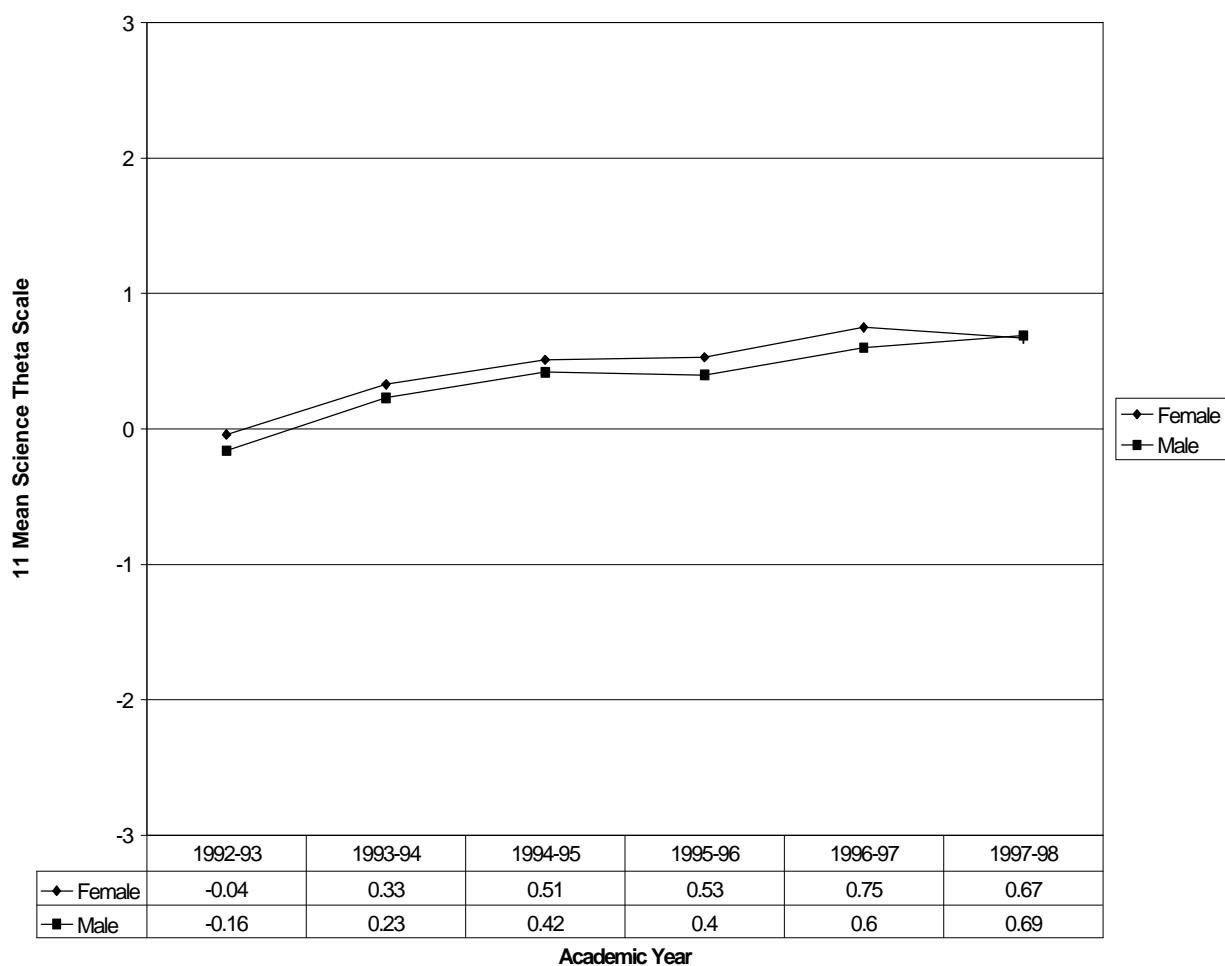
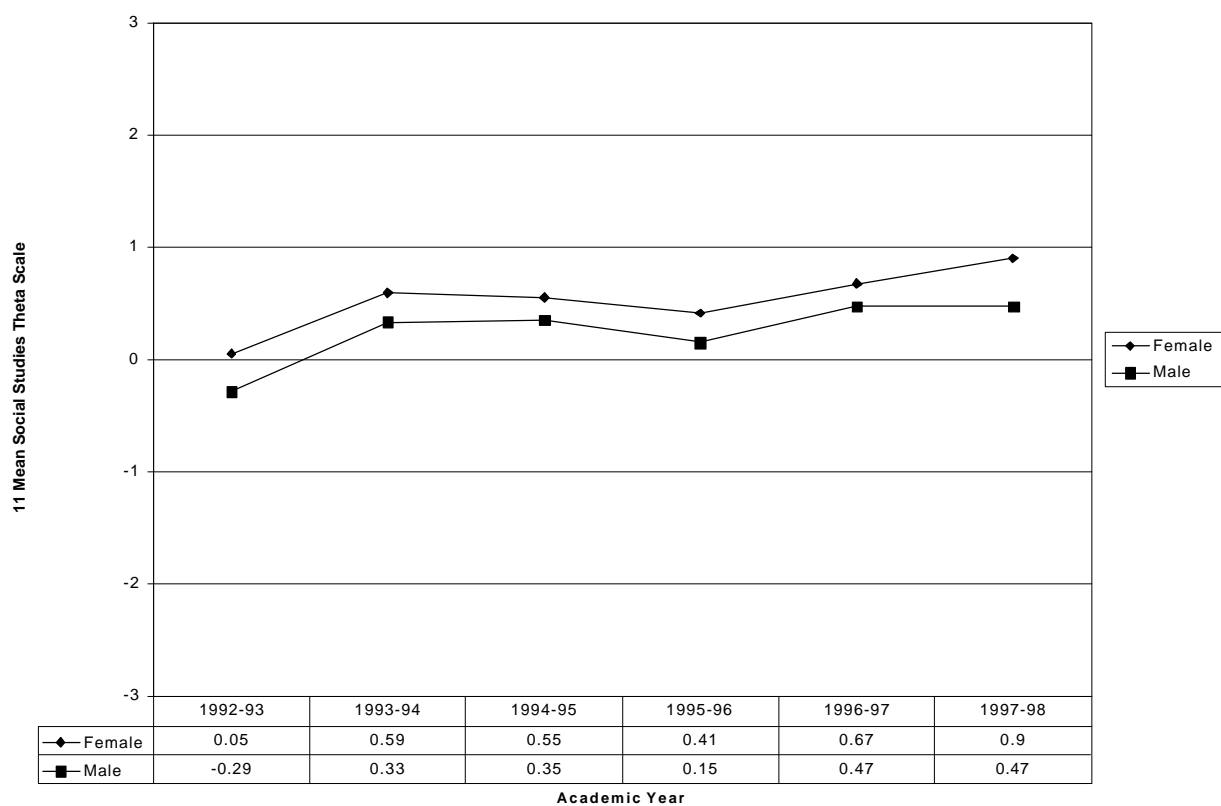


Figure A12. KIRIS Social Studies Theta for Grade 11/12 by Gender, 1993-1998



Appendix B⁴²

KIRIS Cycle 2 and Cycle 3 Subject Area Performance as
Measured by the Theta scale for All Grades Tested by
Race

⁴² Data Tables were added to all Figures in Appendix B by KDE.

Figure B1. KIRIS Reading Theta for Grade 4 by Race, 1993-1998

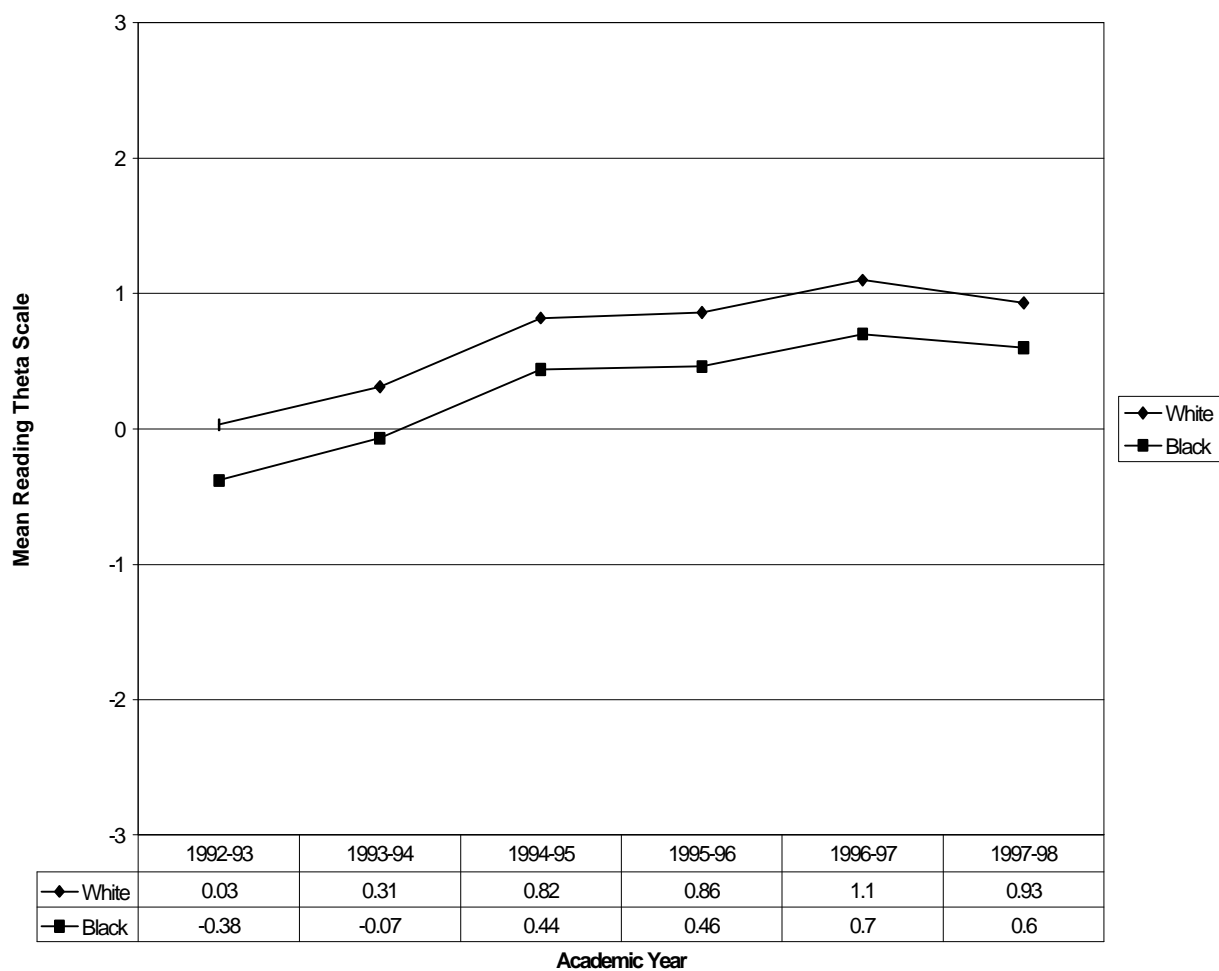


Figure B2. KIRIS Mathematics Theta for Grade 4/5 by Race, 1993-1998

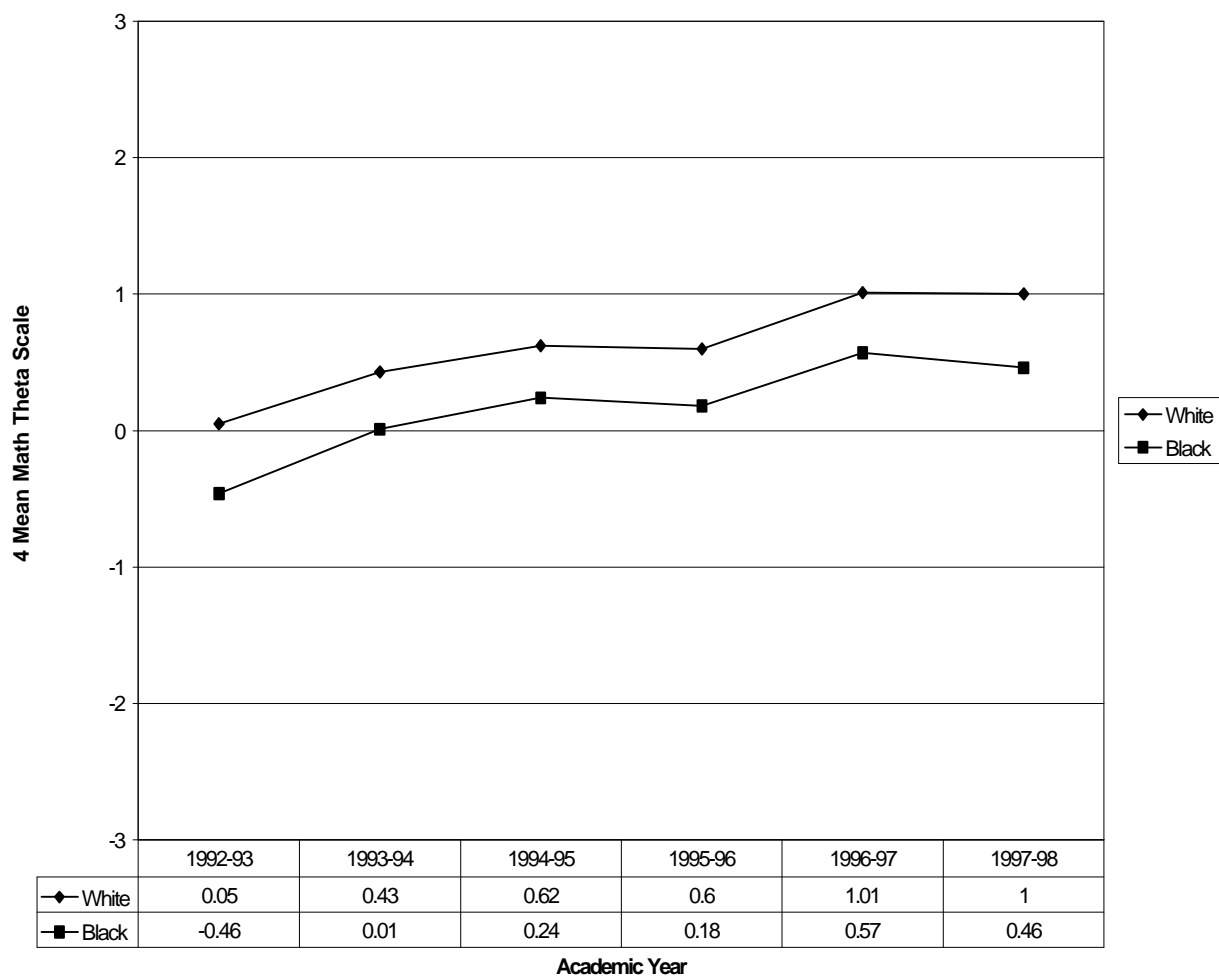


Figure B3. KIRIS Science Theta for Grade 4 by Race, 1993-1998

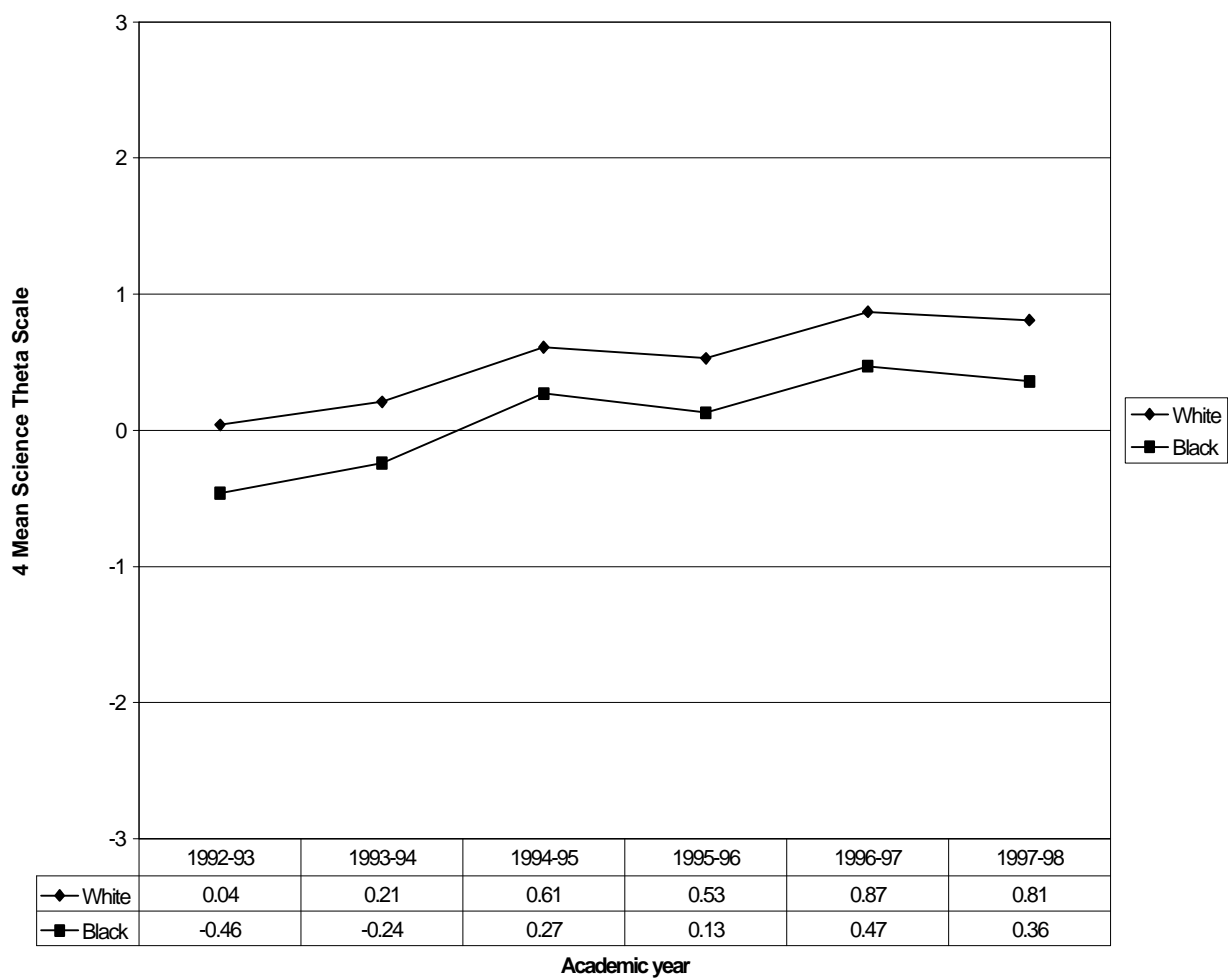


Figure B4. KIRIS Social Studies Theta for Grade 4/5 by Race, 1993-1998

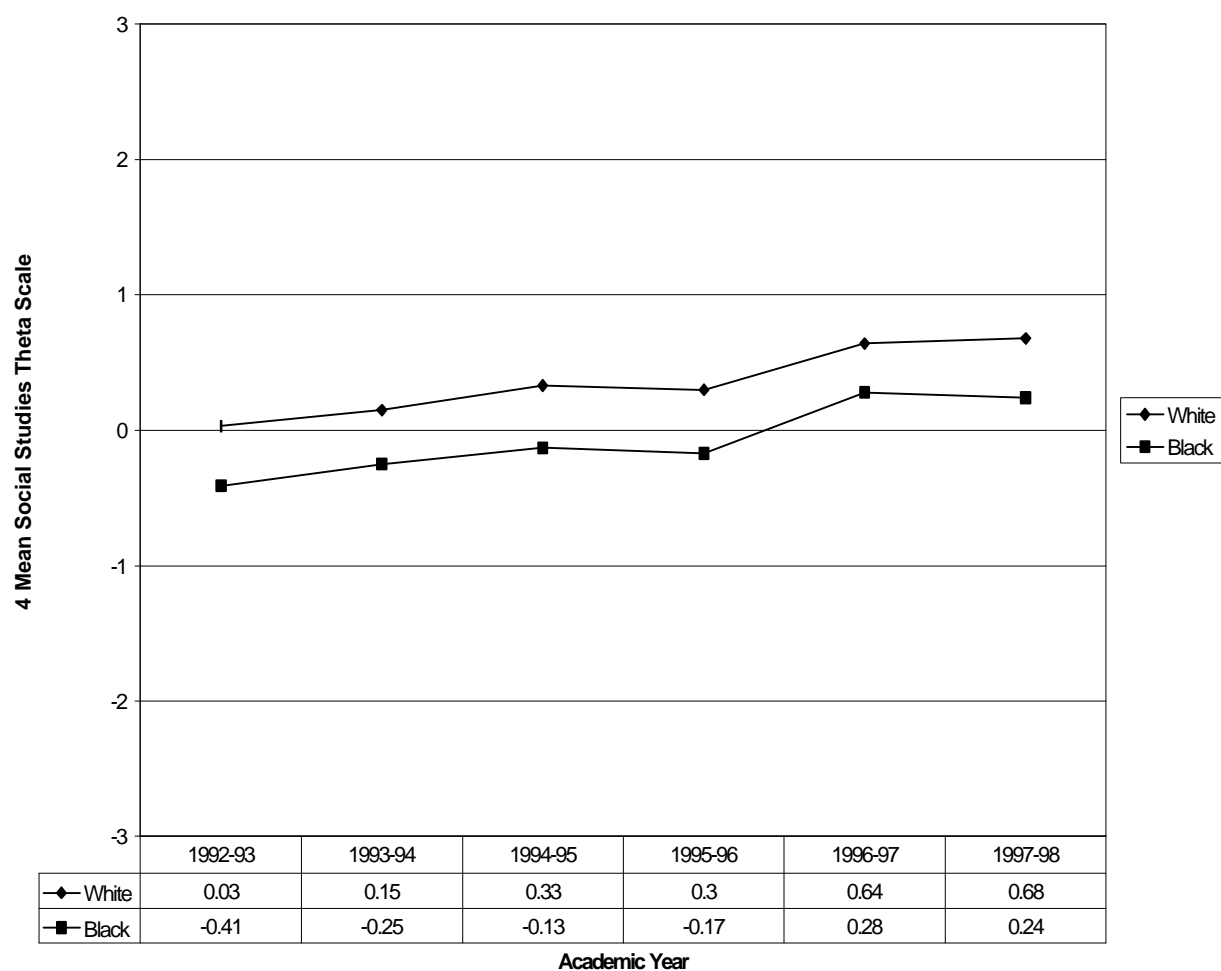


Figure B5. KIRIS Reading Theta for Grade 7/8 by Race, 1993-1998

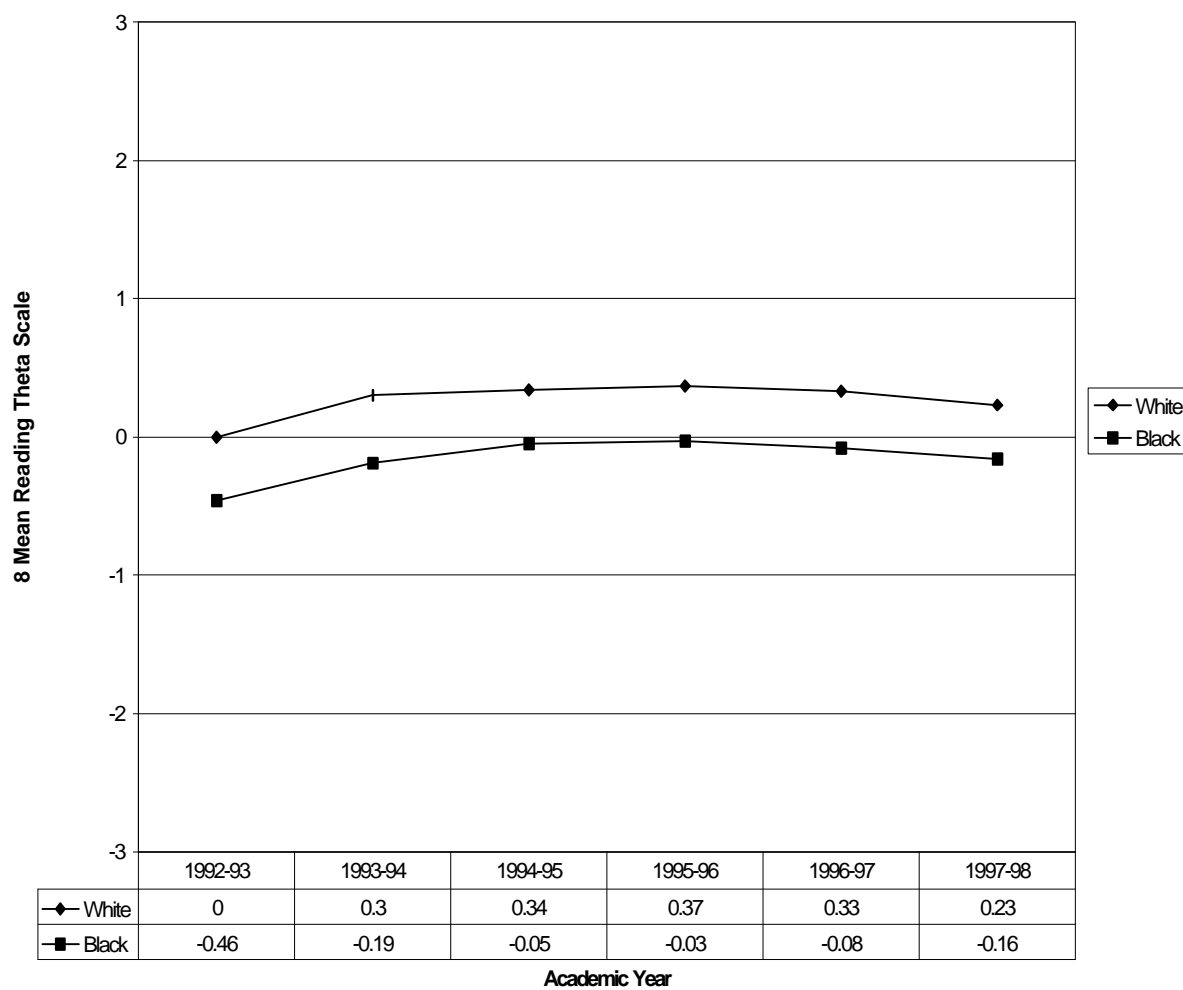


Figure B6. KIRIS Mathematics Theta for Grade 8 by Race, 1993-1998

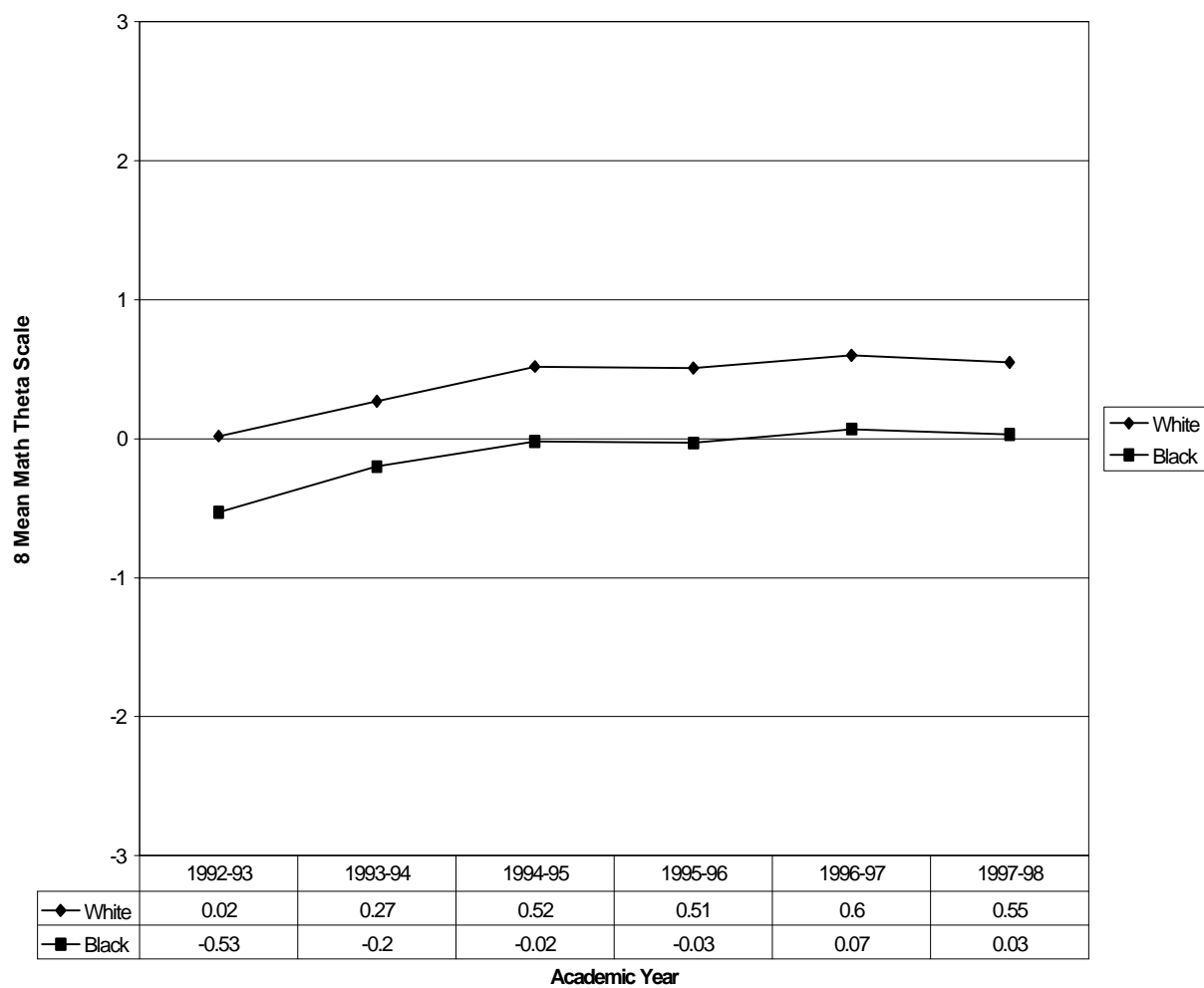


Figure B7. KIRIS Science Theta for Grade 7/8 by Race, 1993-1998

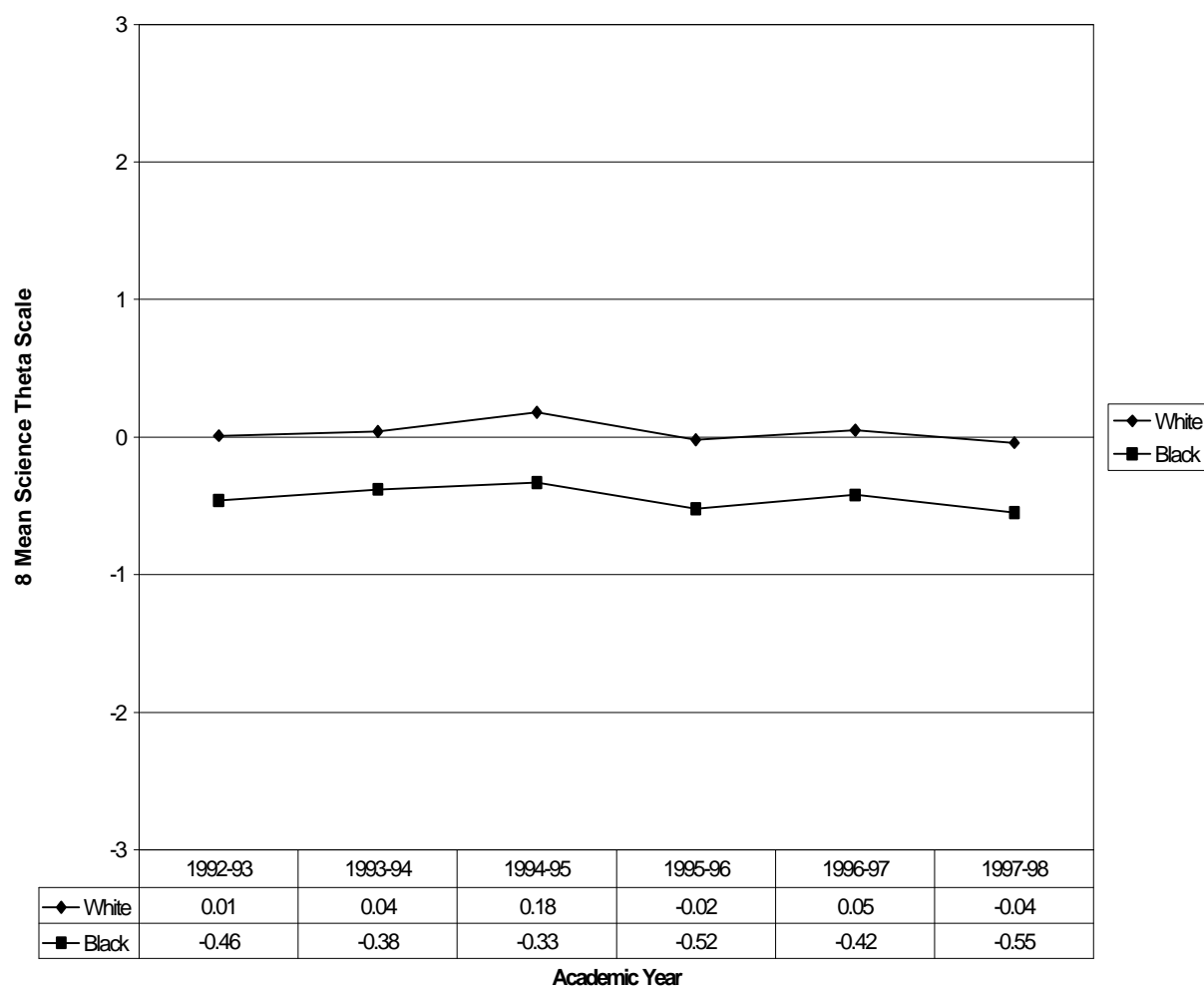


Figure B8. KIRIS Social Studies Theta for Grade 8 by Race, 1993-1998

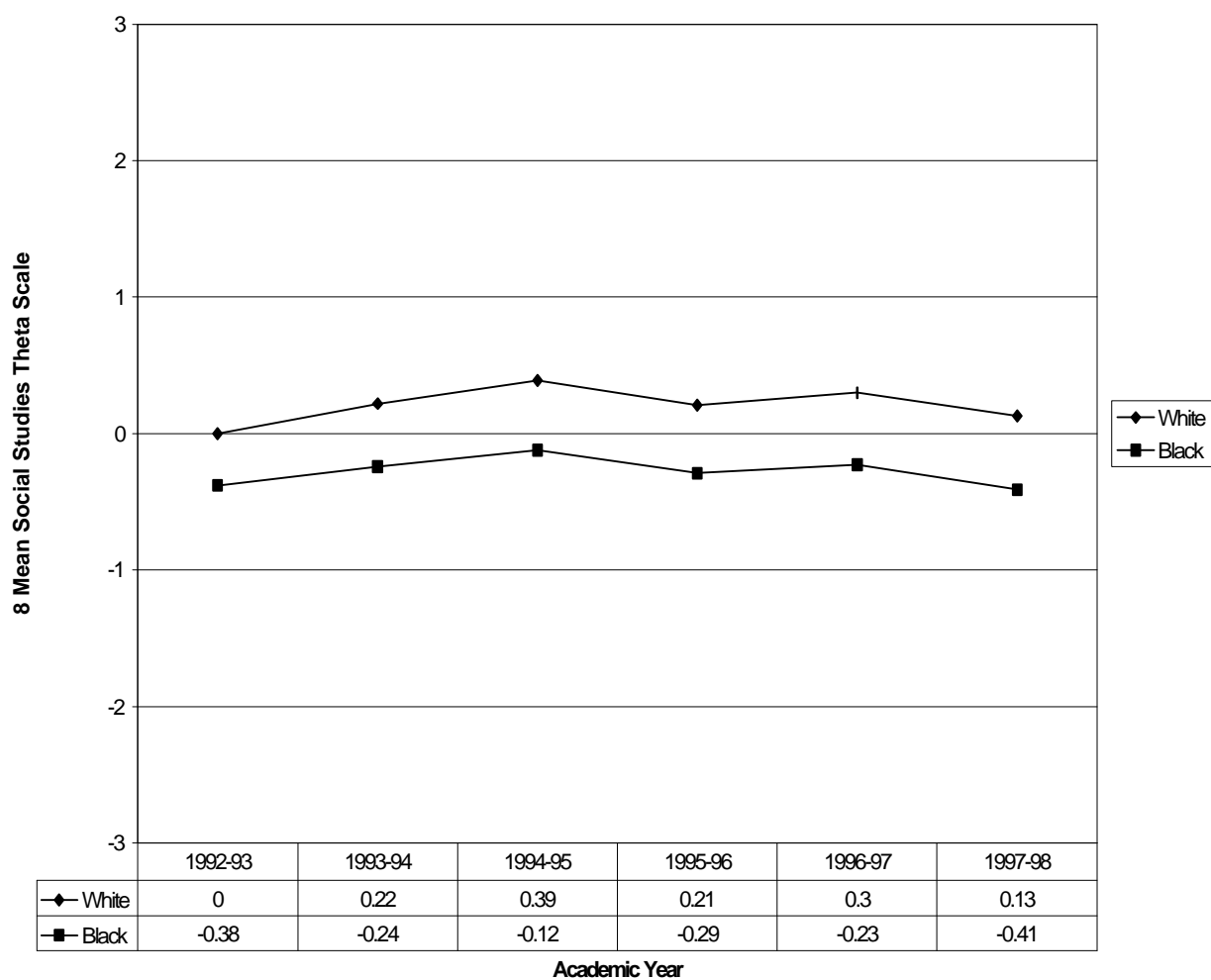


Figure B9. KIRIS Reading Theta for Grade 11/12 by Race, 1993-1998

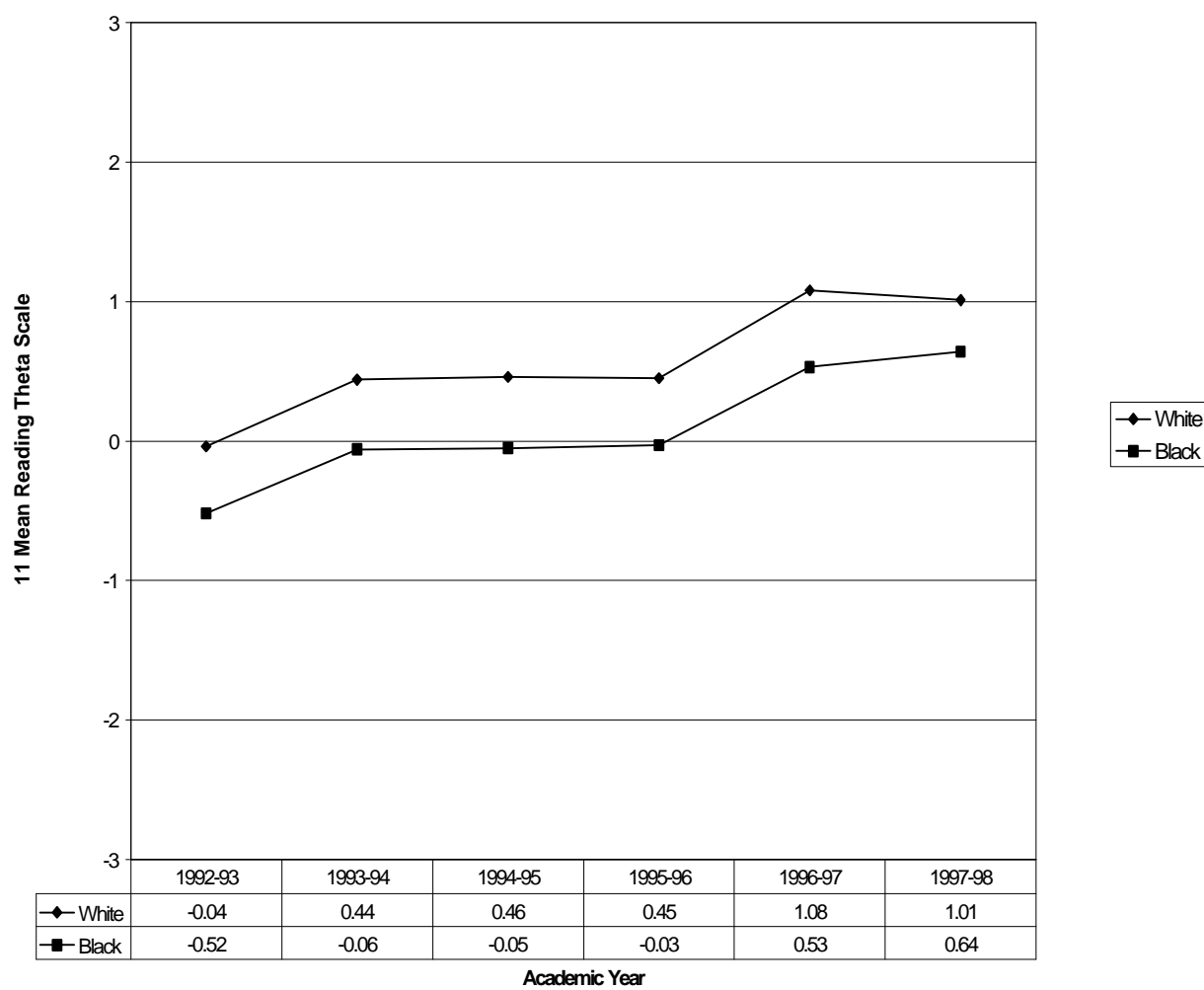


Figure B10. KIRIS Mathematics Theta for Grade 11/12 by Race, 1993-1998

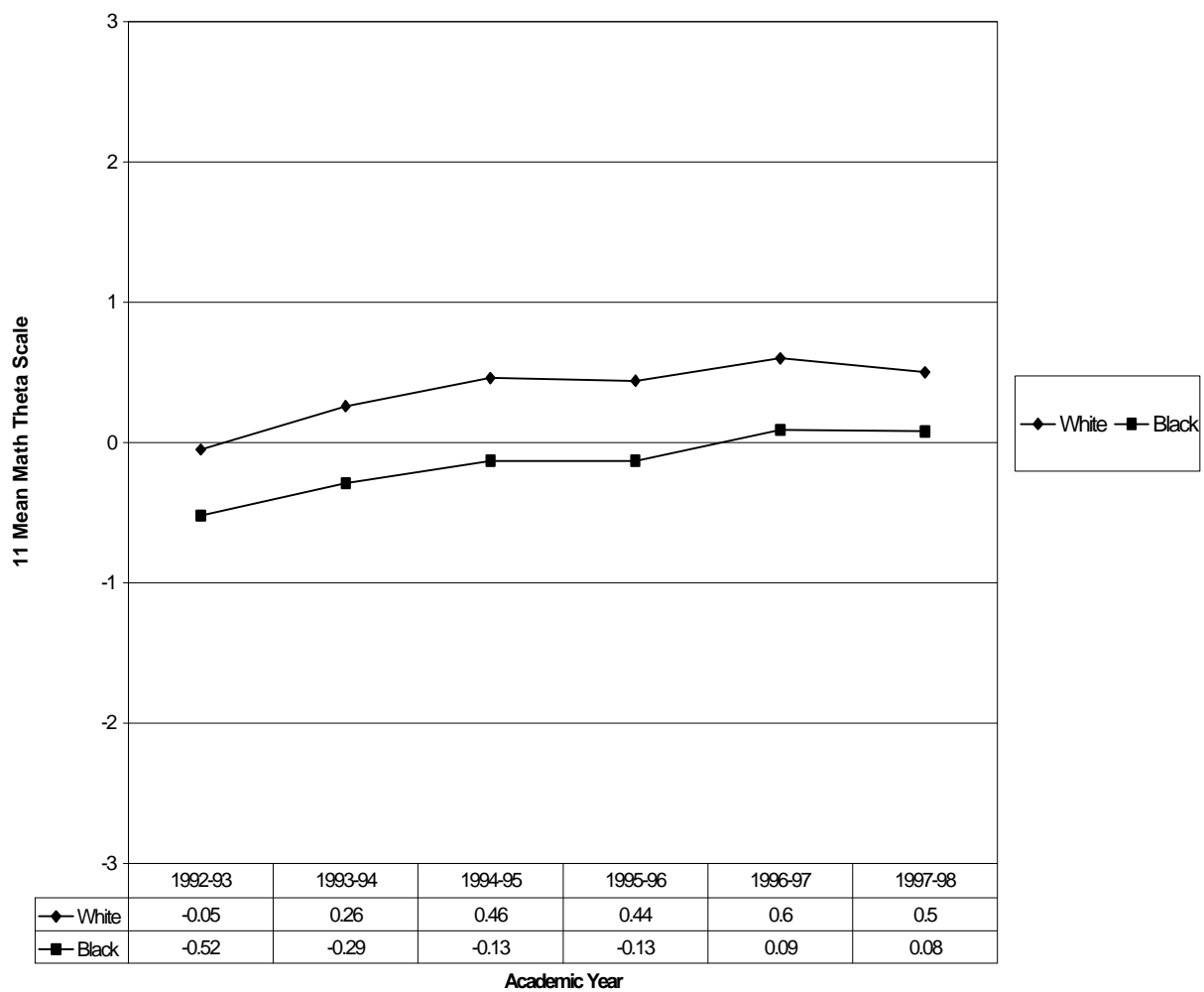


Figure B11. KIRIS Science Theta for Grade 11/12 by Race, 1993-1998

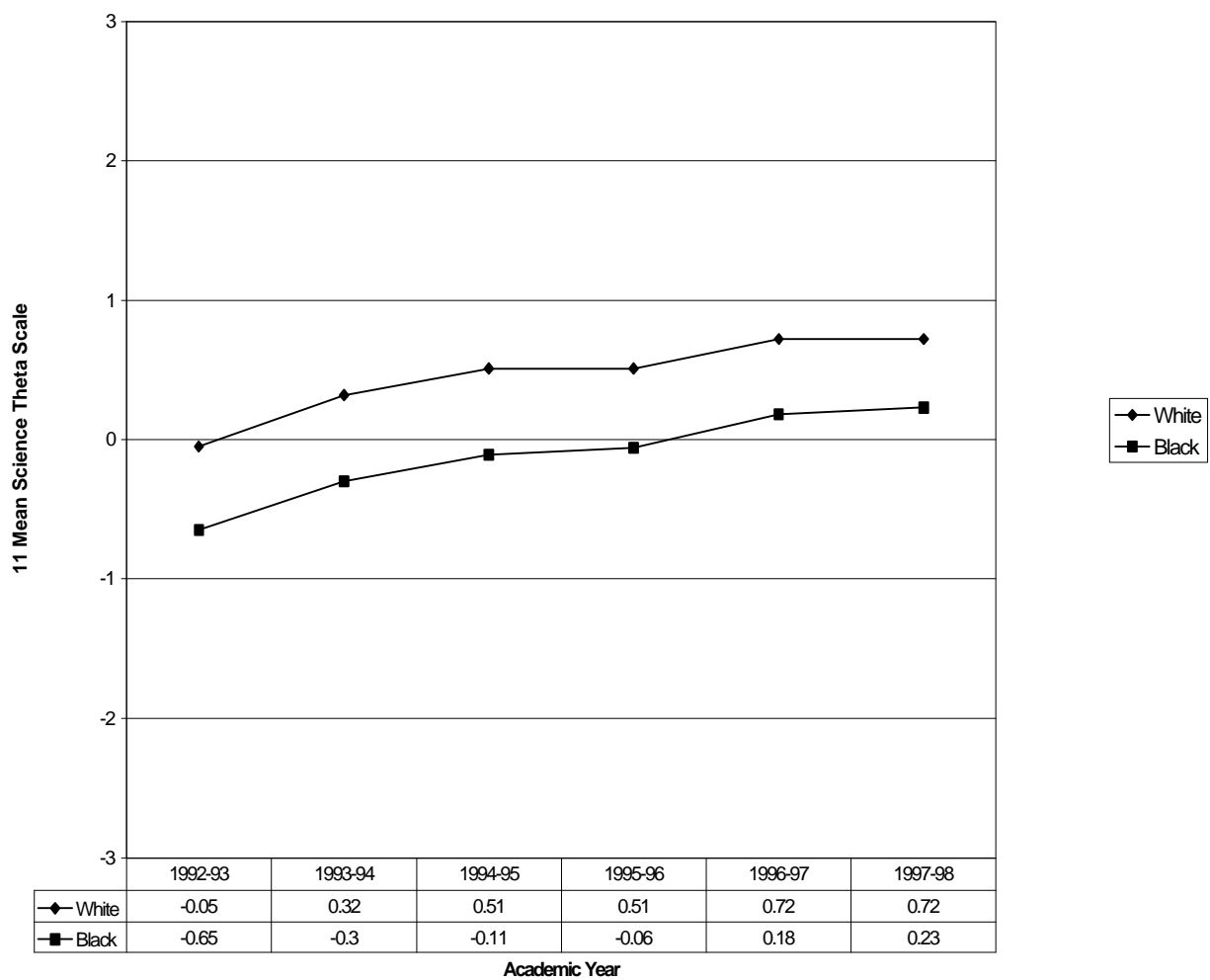
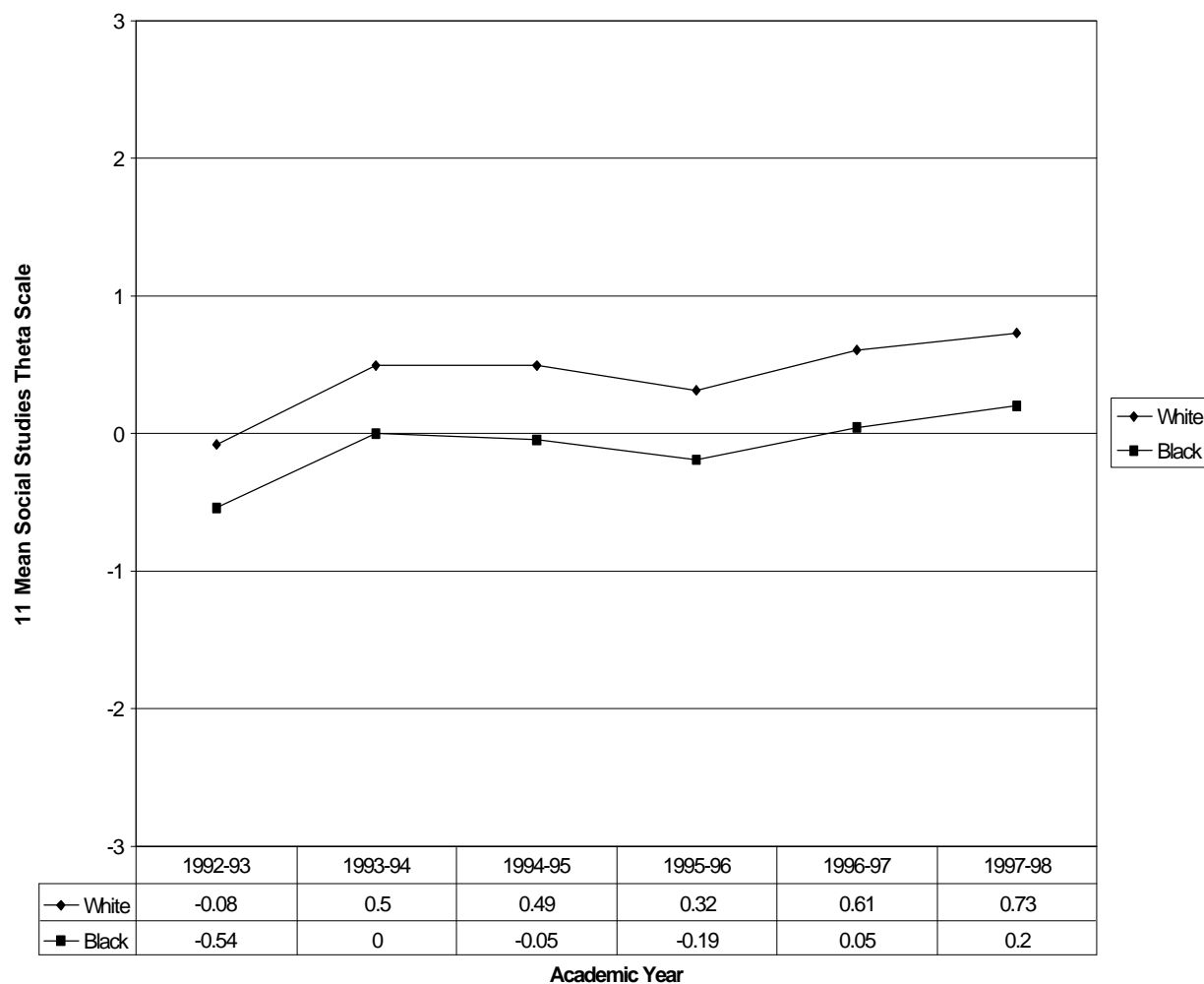


Figure B12. KIRIS Social Studies Theta for Grade 11/12 by Race, 1993-1998



Appendix C⁴³

KIRIS Cycle 2 and Cycle 3 Subject Area Performance as
Measured by the Constructed Response Index for All
Grades Tested by Gender and Race

⁴³ Data Tables added to all Figures in Appendix C by KDE.

Figure C1. KIRIS Reading Constructed Response Index Scores for Grade 4 by Gender and Race, 1993-1998

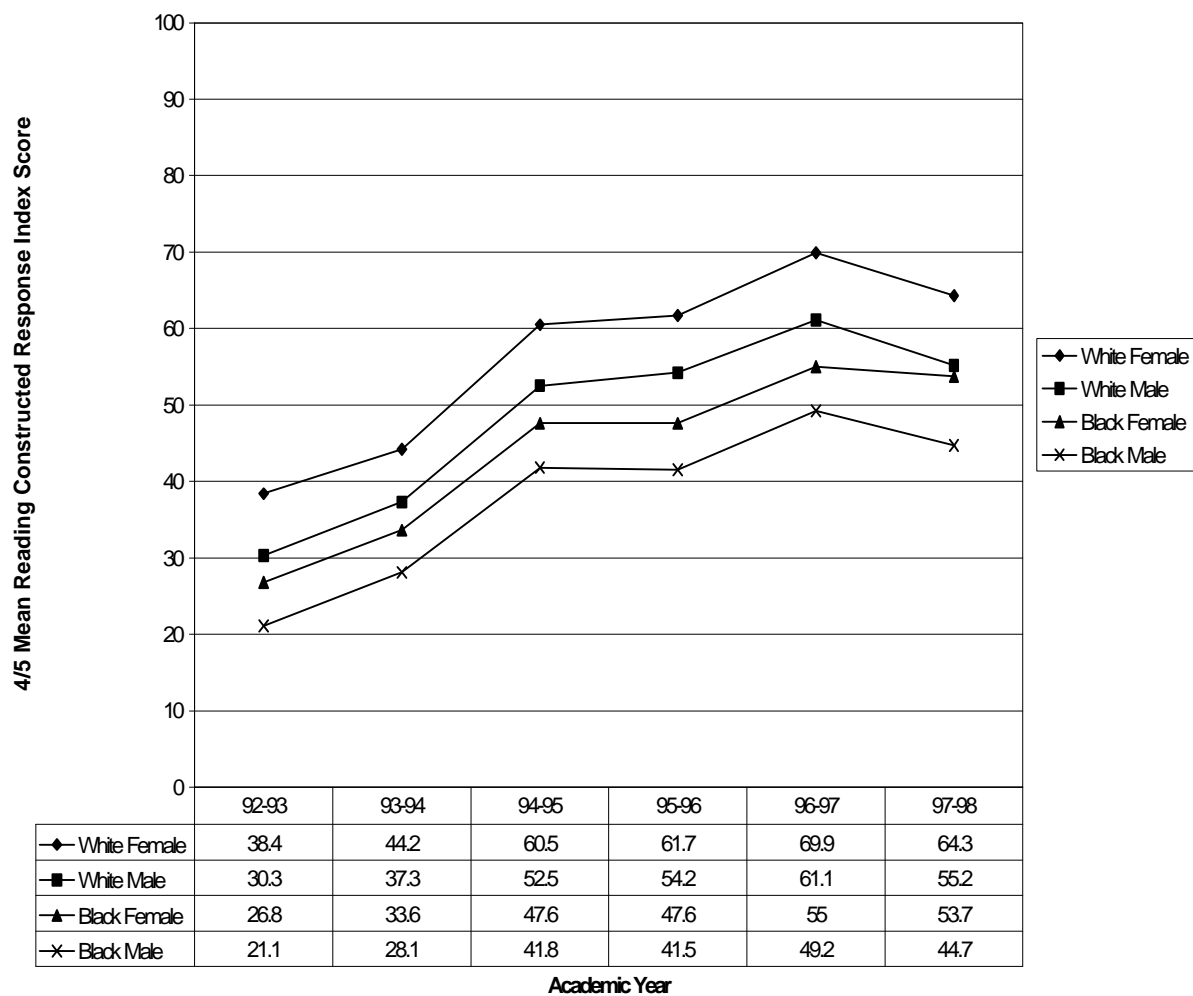


Figure C2. KIRIS Mathematics Constructed Response Index Scores for Grade 4/5 by Gender and Race, 1993-1998

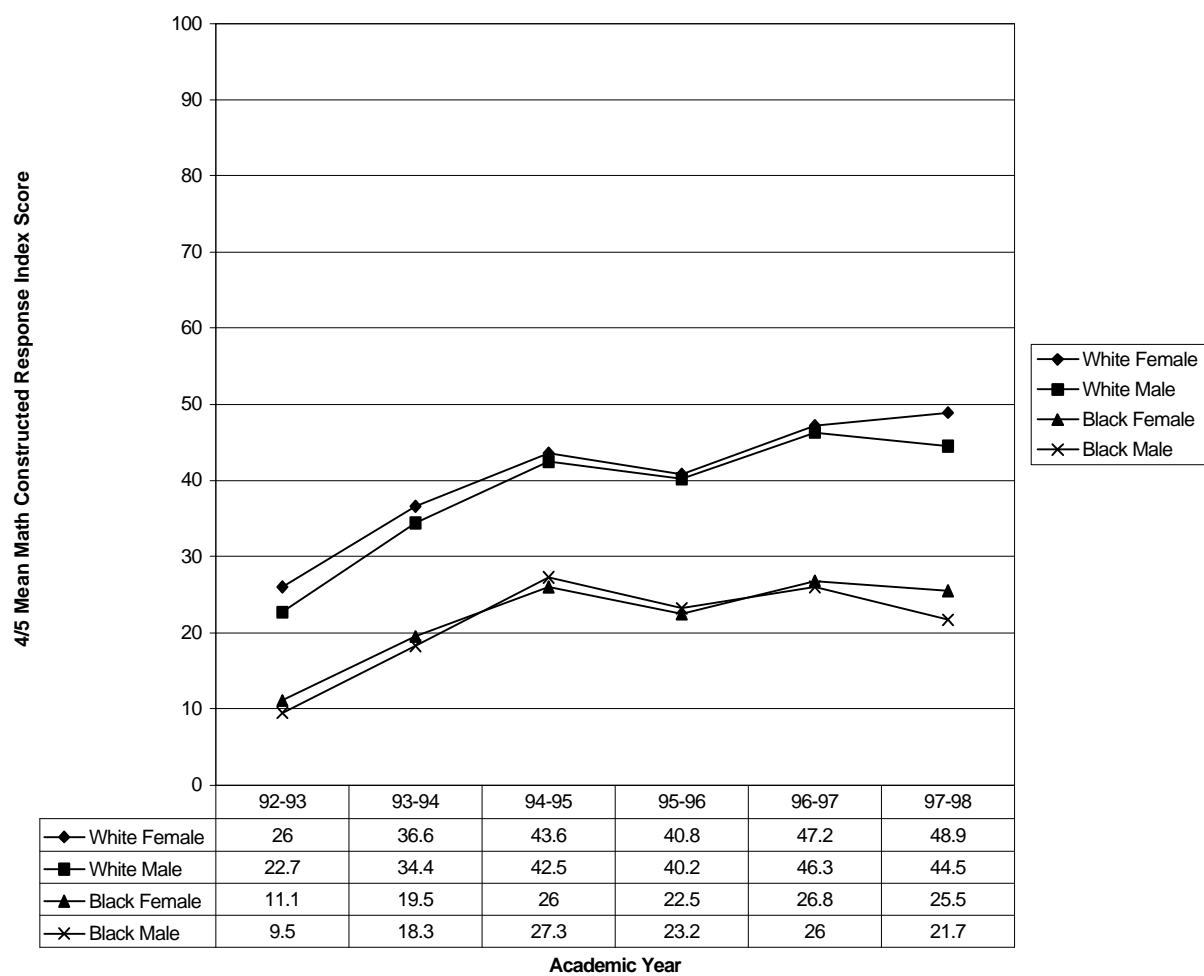


Figure C3. KIRIS Science Constructed Response Index Scores for Grade 4 by Gender and Race, 1993-1998

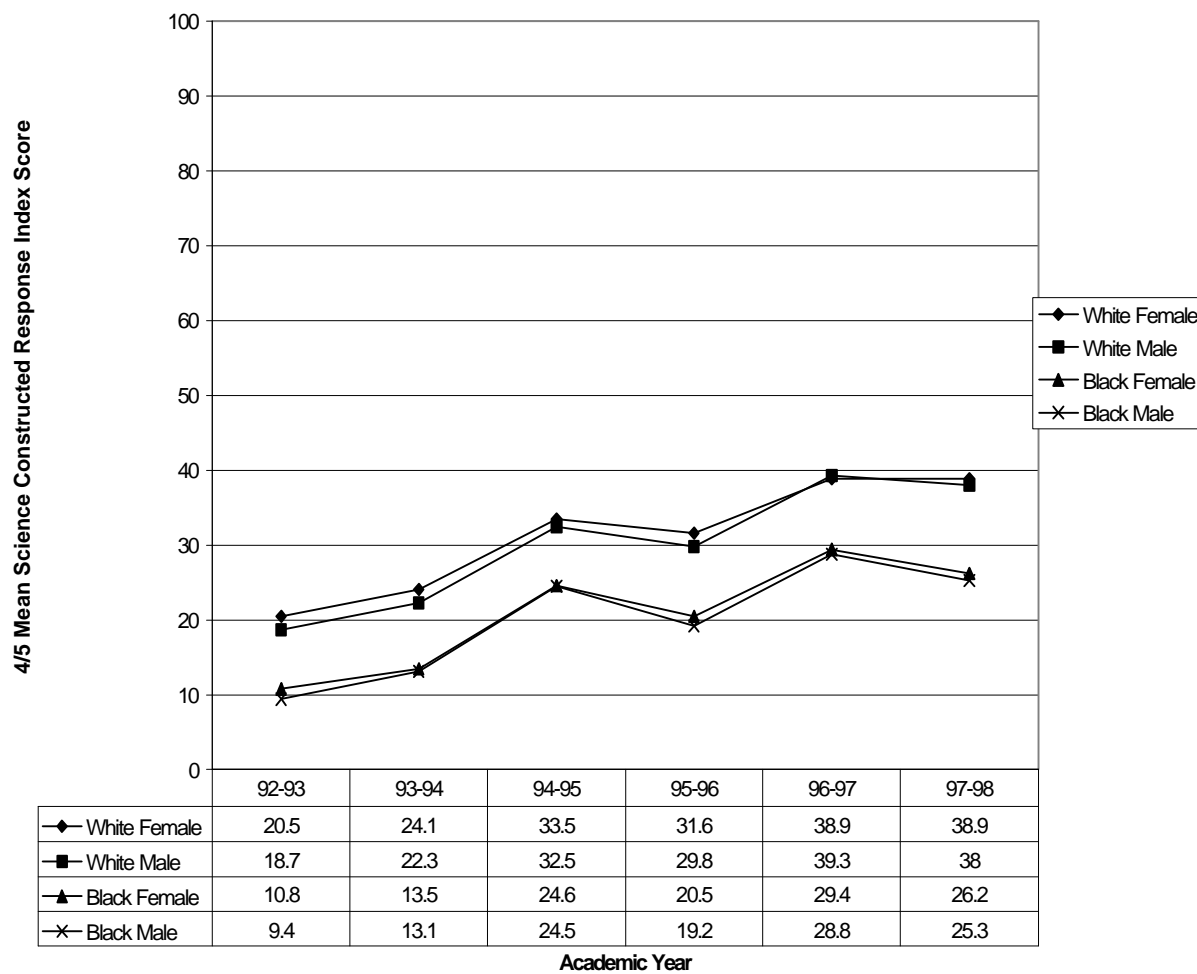


Figure C4. KIRIS Social Studies Constructed Response Index Score for Grade 4/5 by Gender and Race, 1993-1998

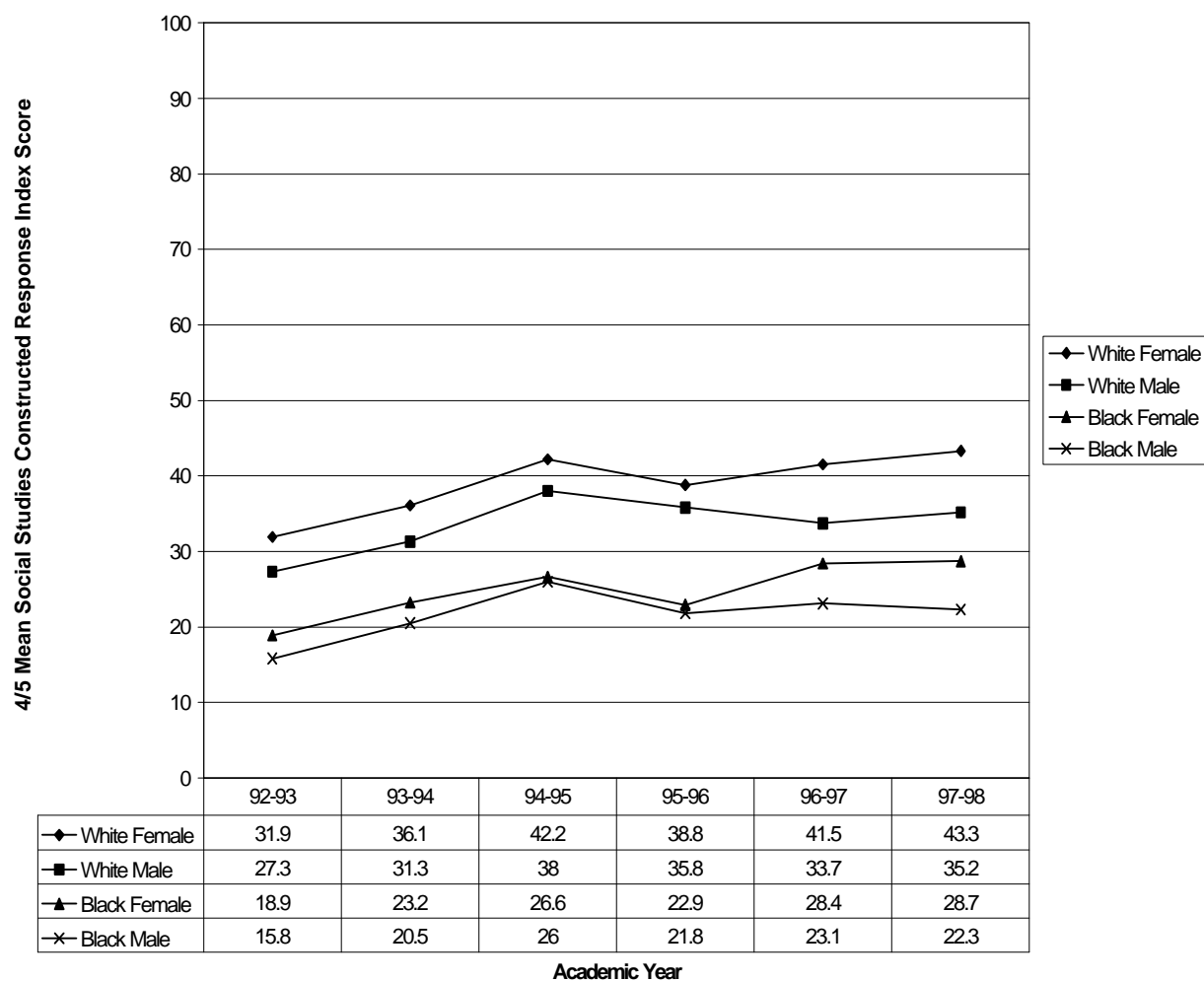


Figure C5. KIRIS Reading Constructed Response Index Score for Grade 7/8 by Gender and Race, 1993-1998

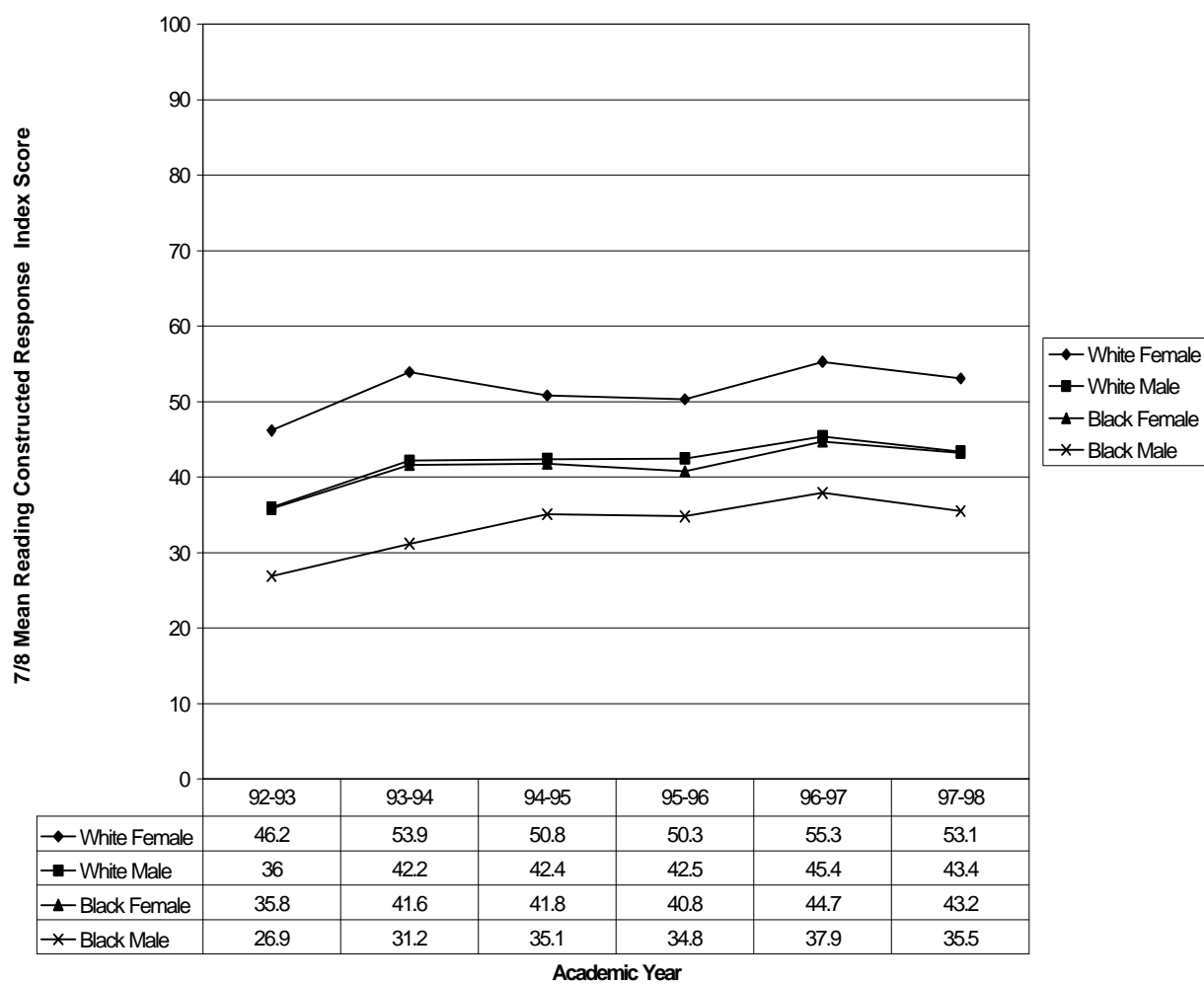


Figure C6. KIRIS Mathematics Constructed Response Index Score for Grade 8 by Gender and Race, 1993-1998

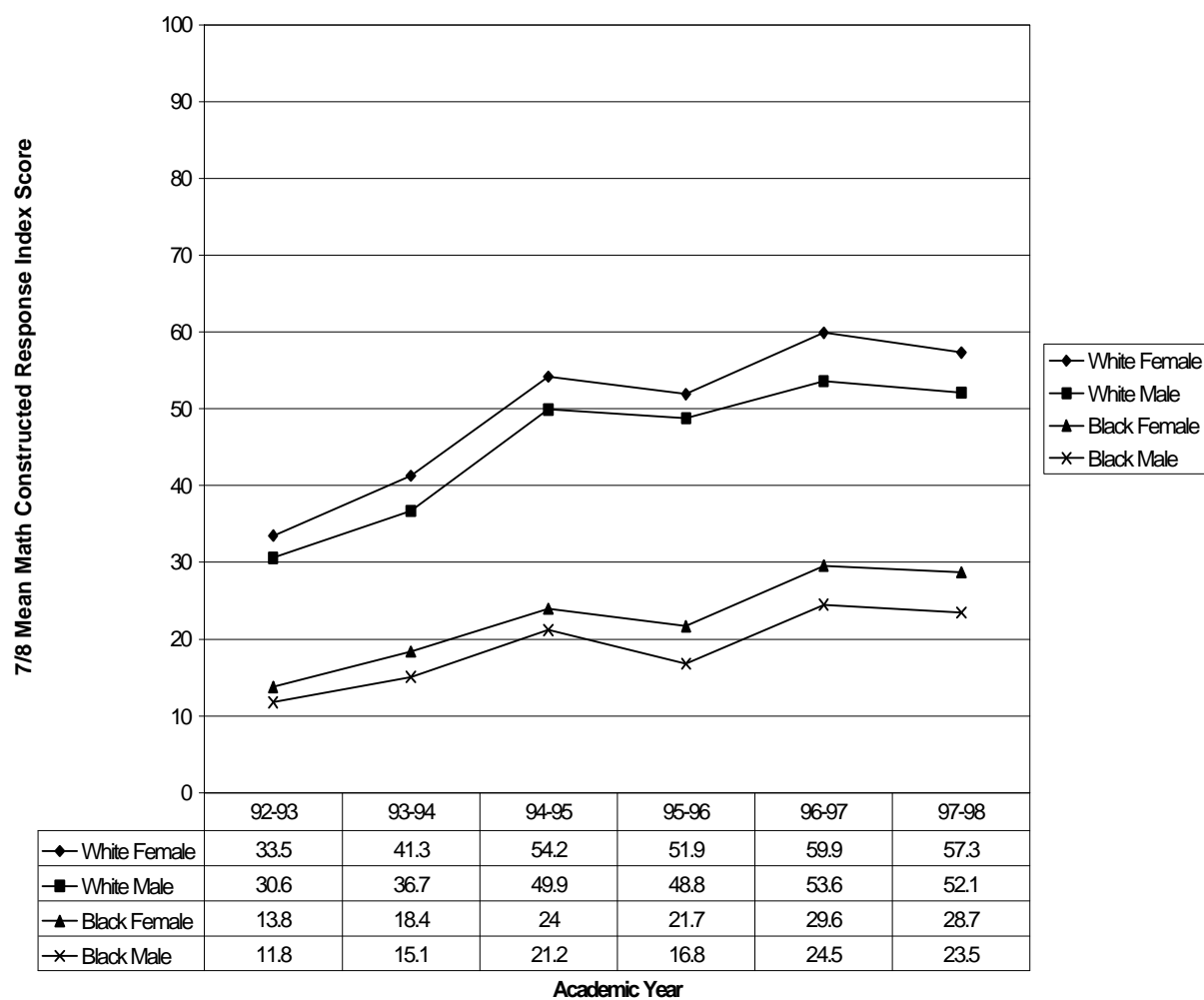


Figure C7. KIRIS Science Constructed Response Index Score for Grade 7/8 by Gender and Race, 1993-1998

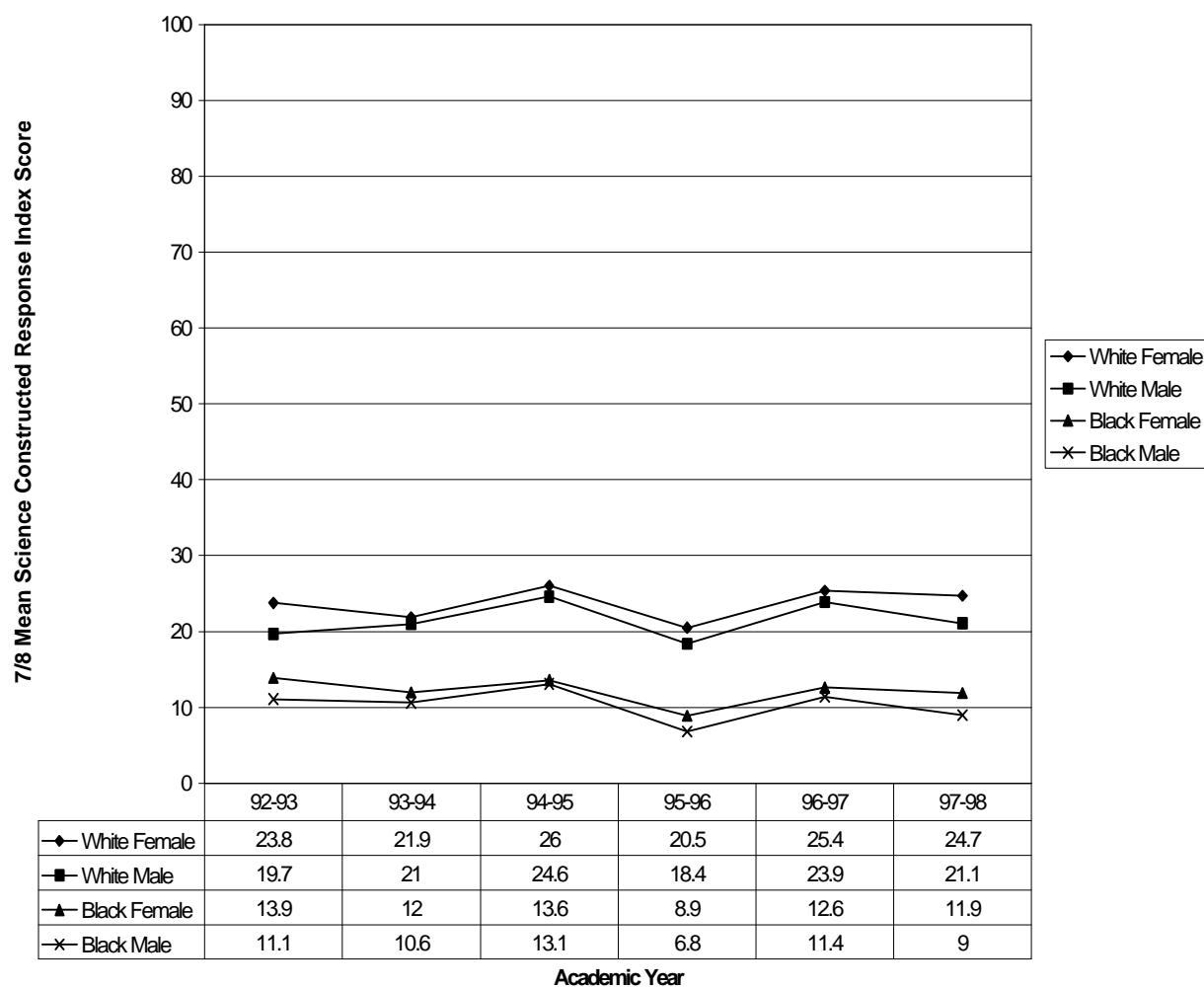


Figure C8. KIRIS Social Studies Constructed Response Index Score for Grade 8 by Gender and Race, 1993-1998

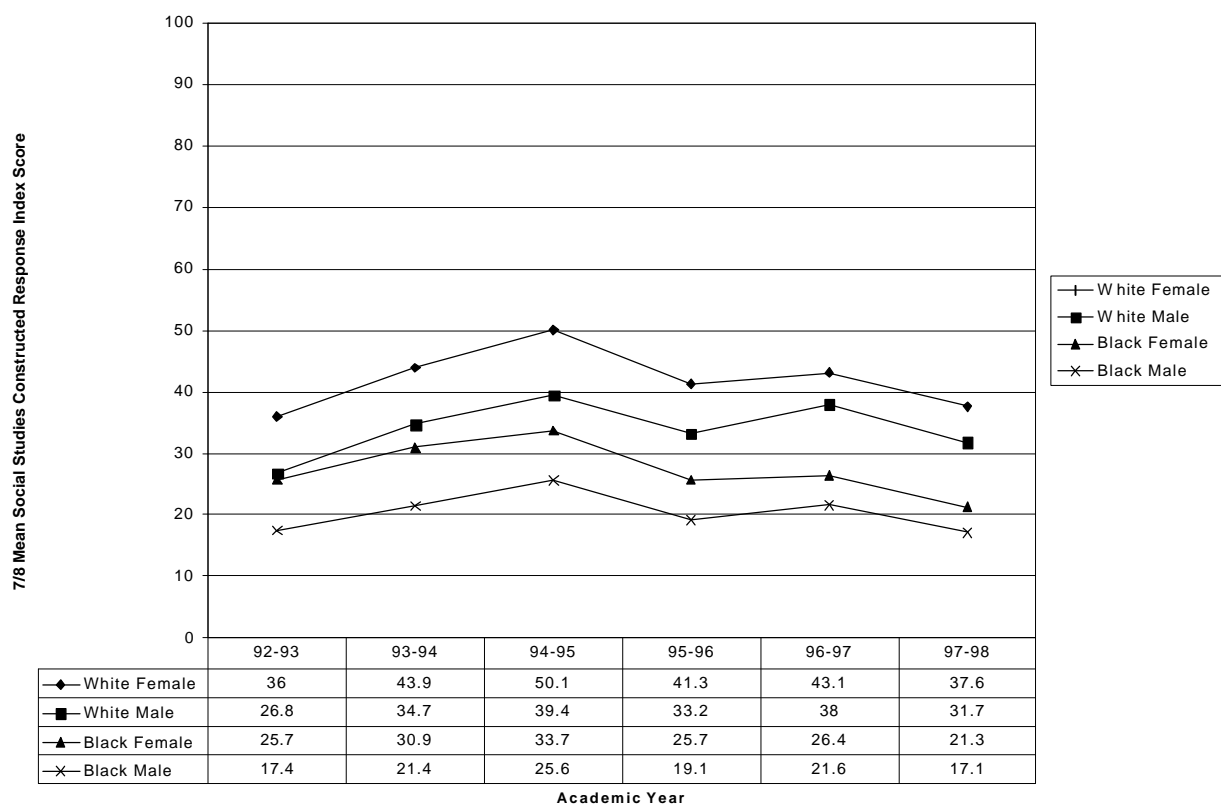


Figure C9. KIRIS Reading Constructed Response Index Scores for Grade 11/12 by Gender and Race, 1993-1998

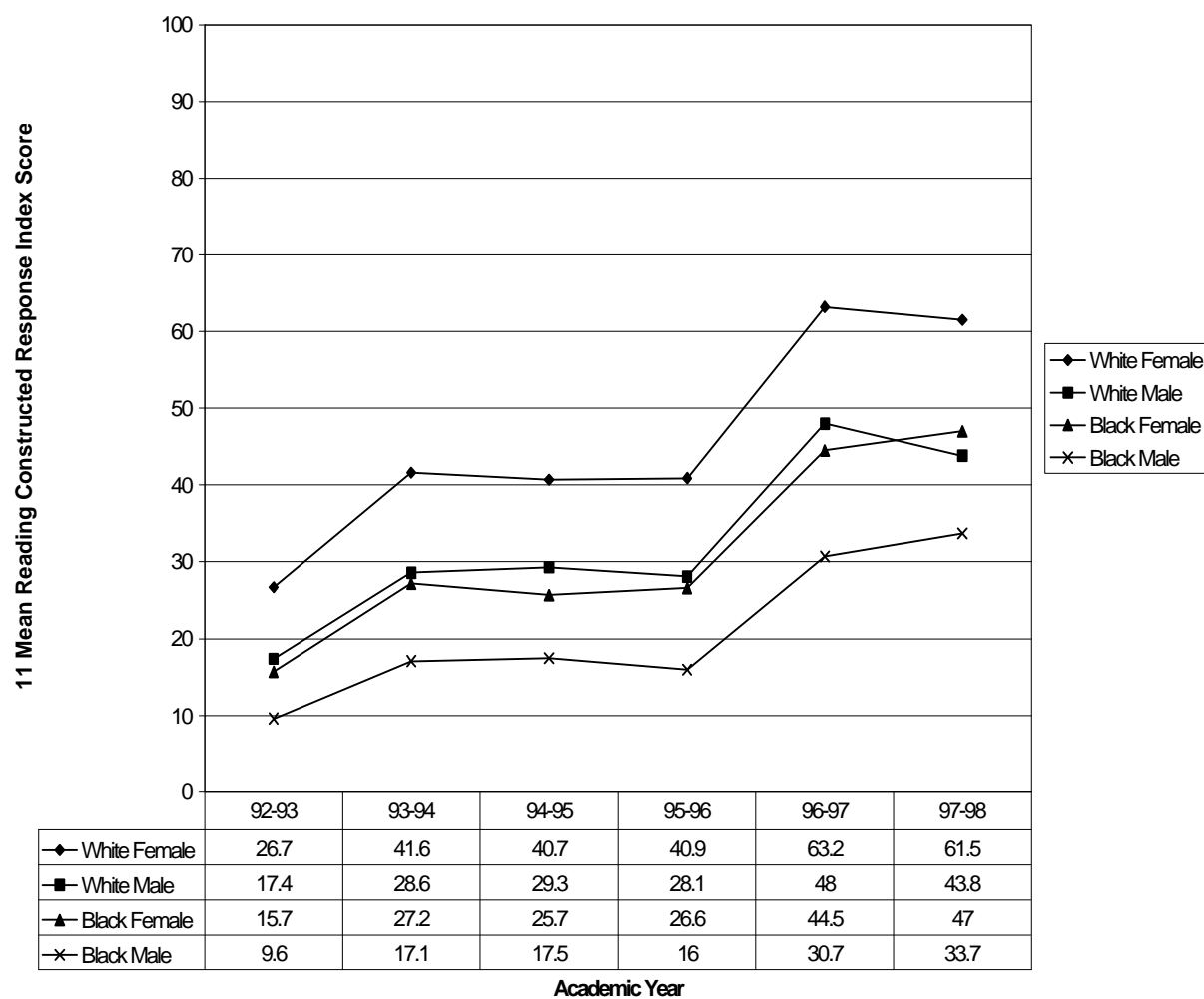


Figure C10. KIRIS Mathematics Constructed Response Index Scores for Grade 11/12 by Gender and Race, 1993-1998

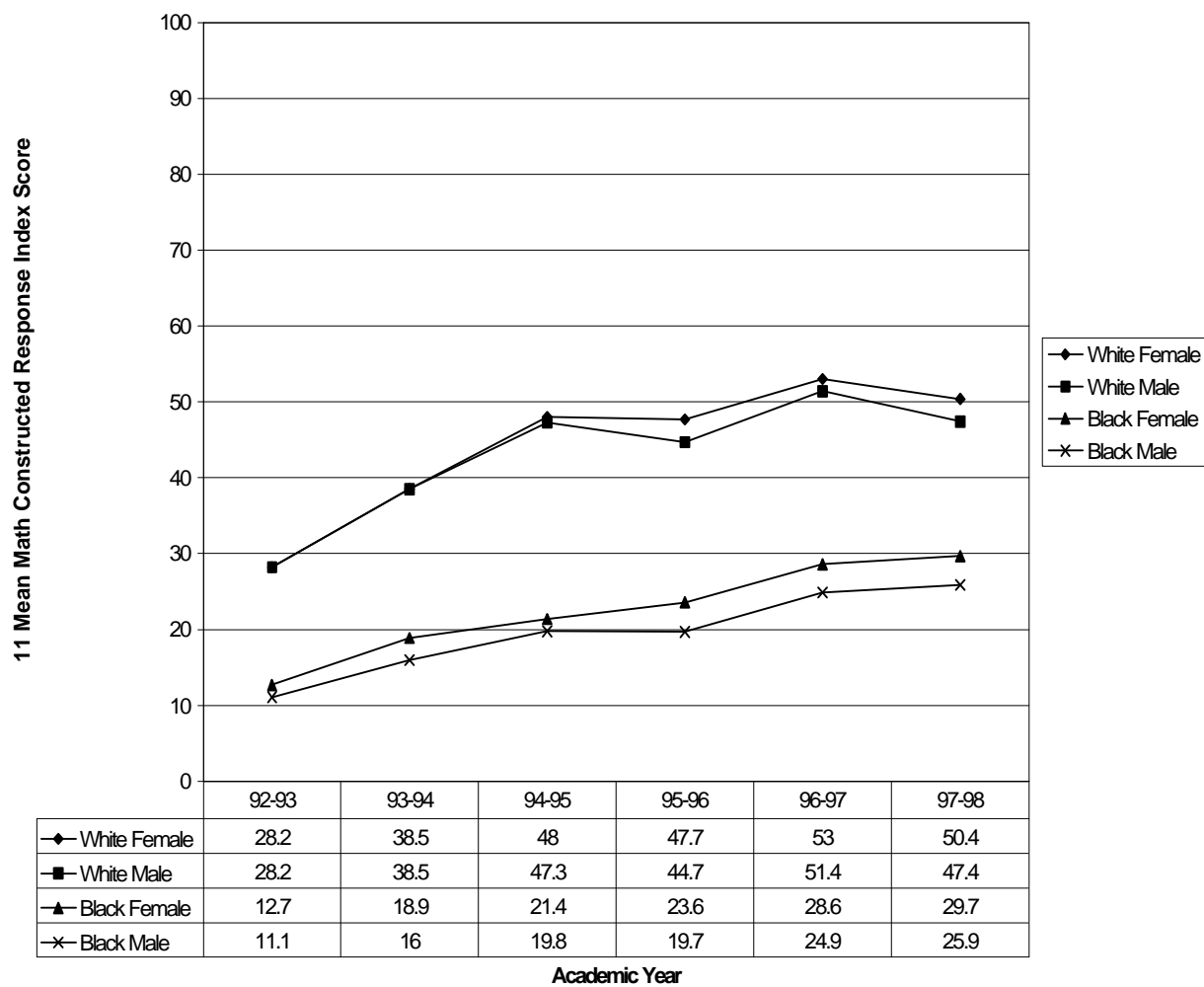


Figure C11. KIRIS Science Constructed Response Index Scores for Grade 11/12 by Gender and Race, 1993-1998

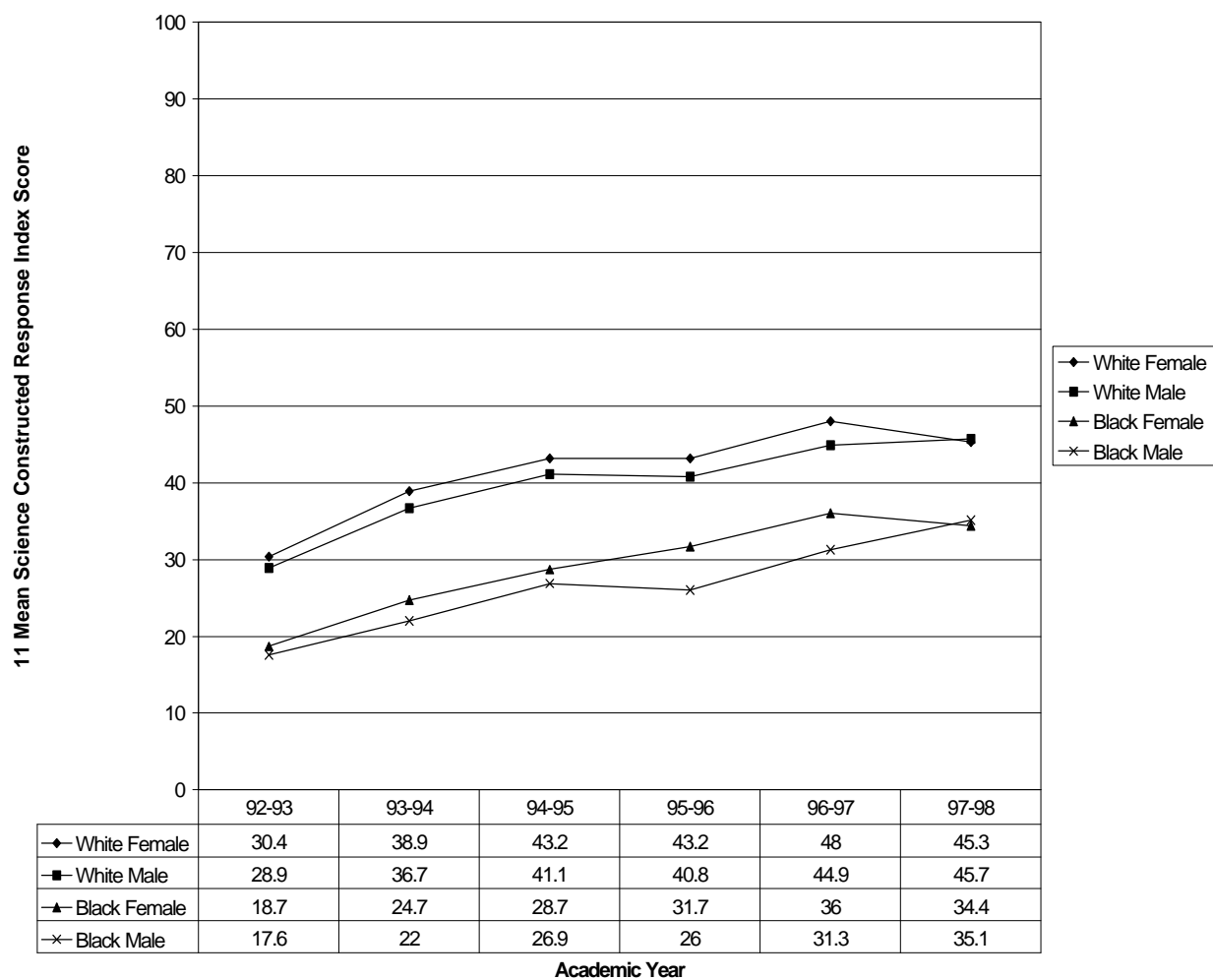
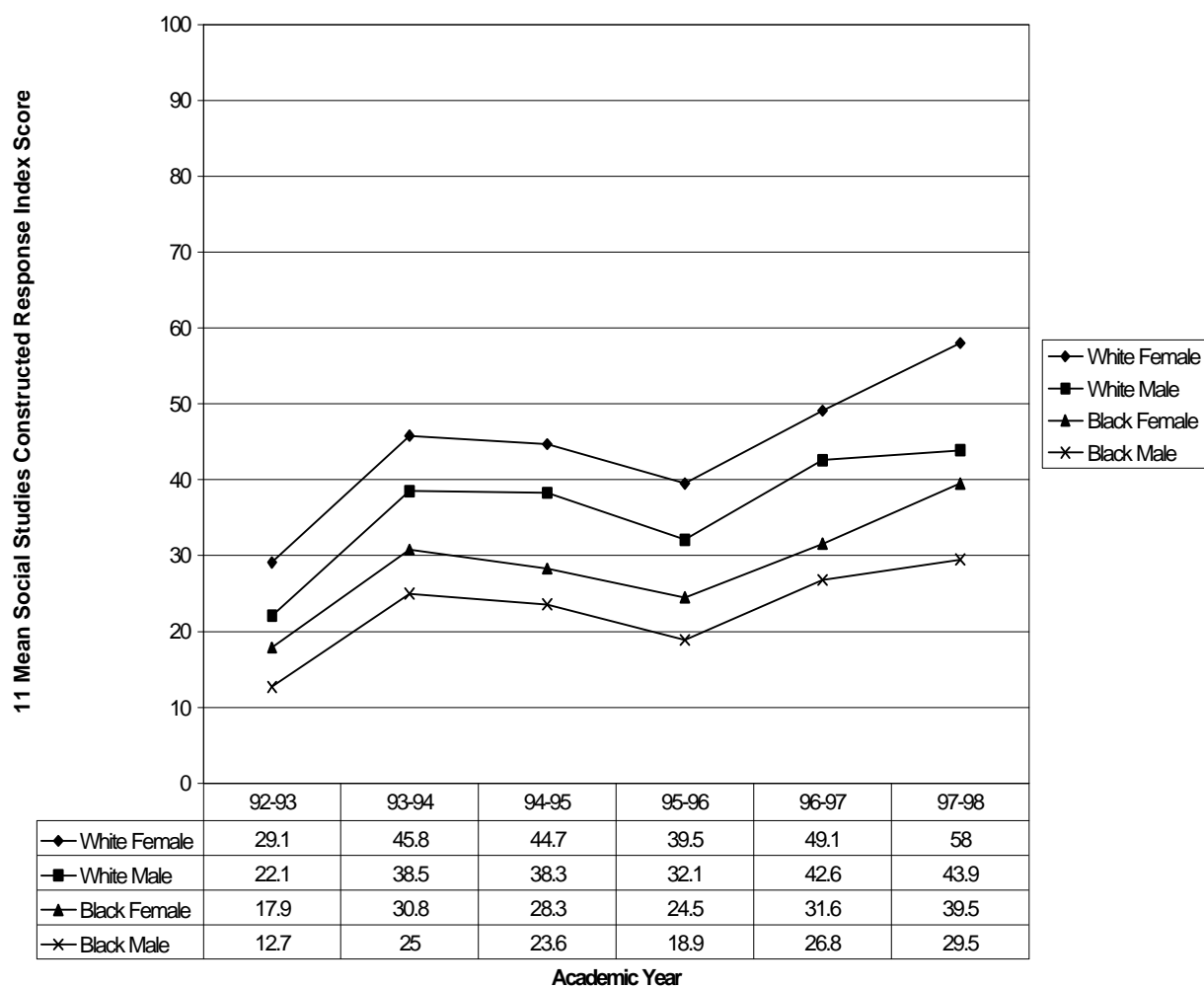


Figure C12. KIRIS Social Studies Constructed Response Index Scores for Grade 11/12 by Gender and Race, 1993-1998



Appendix D⁴⁴

KIRIS Cycle 2 and Cycle 3 Subject Area Performance as
Measured by the Constructed Response Index for All
Grades Tested by Gender

⁴⁴ Data Tables were added for each Figure in Appendix D by KDE.

Figure D1. KIRIS Reading Constructed Response Index Score for Grade 4 by Gender, 1993-1998

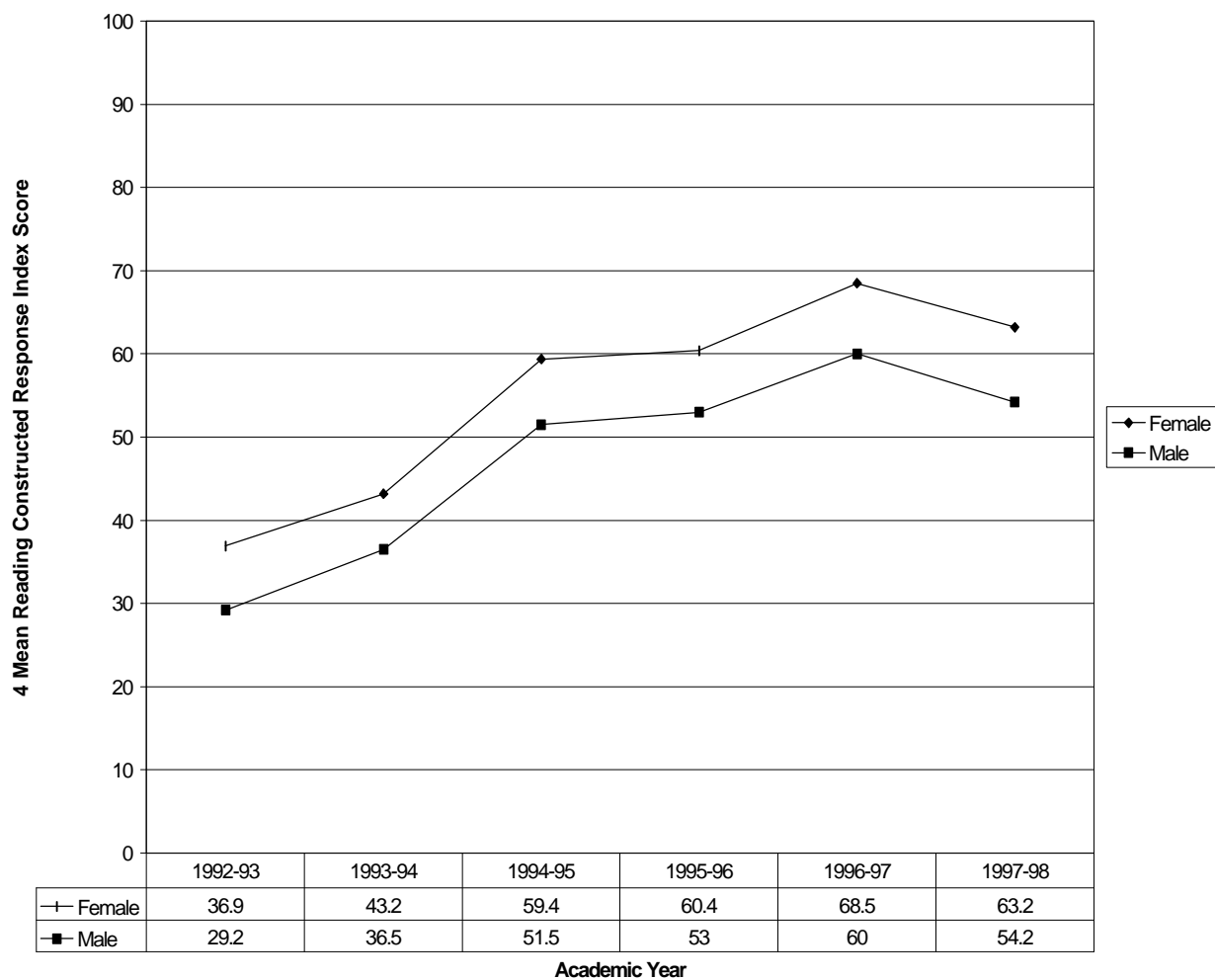


Figure D2. KIRIS Mathematics Constructed Response Index Score for Grade 4/5 by Gender, 1993-1998

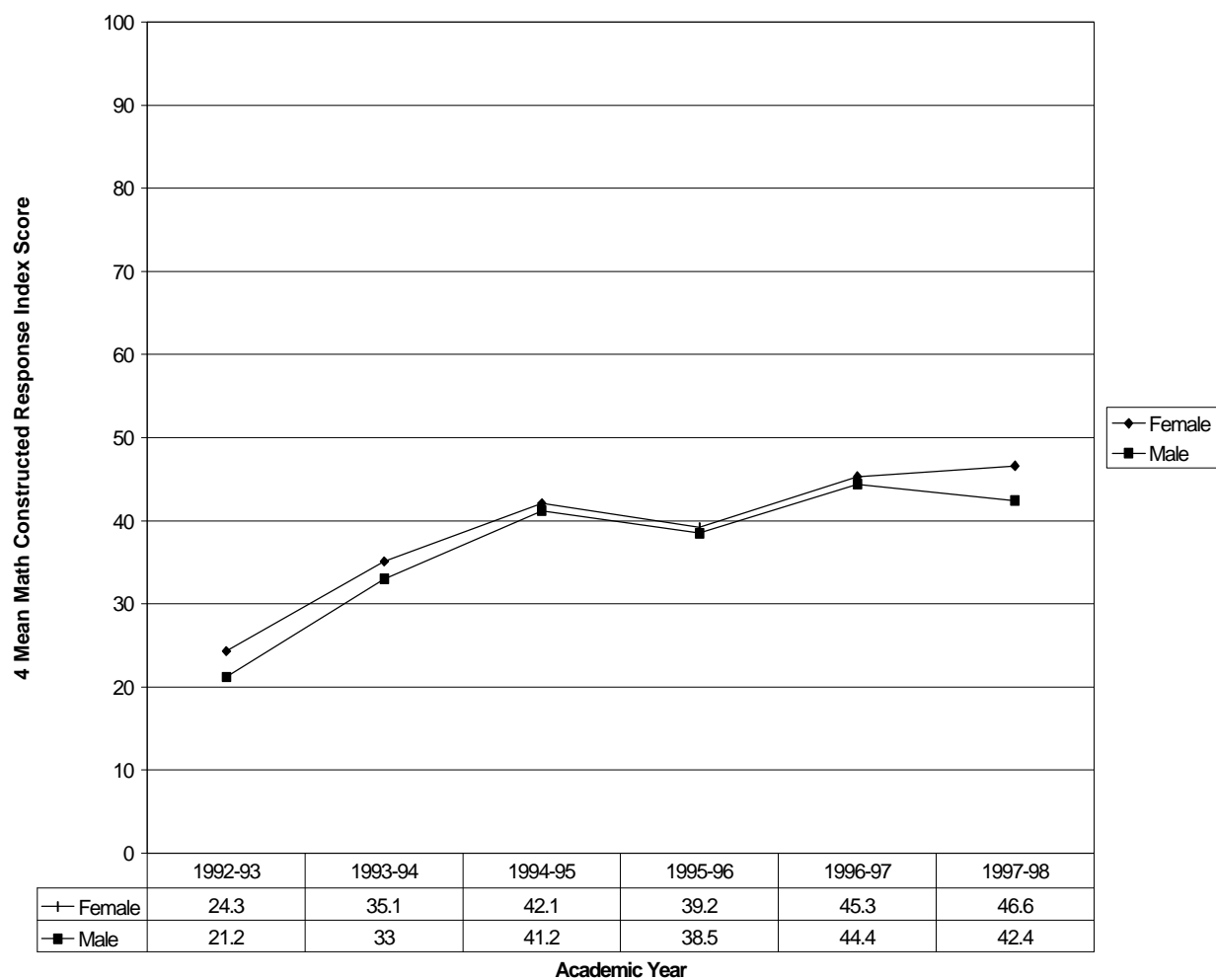


Figure D3. KIRIS Science Constructed Response Index Scores for Grade 4 by Gender, 1993-1998

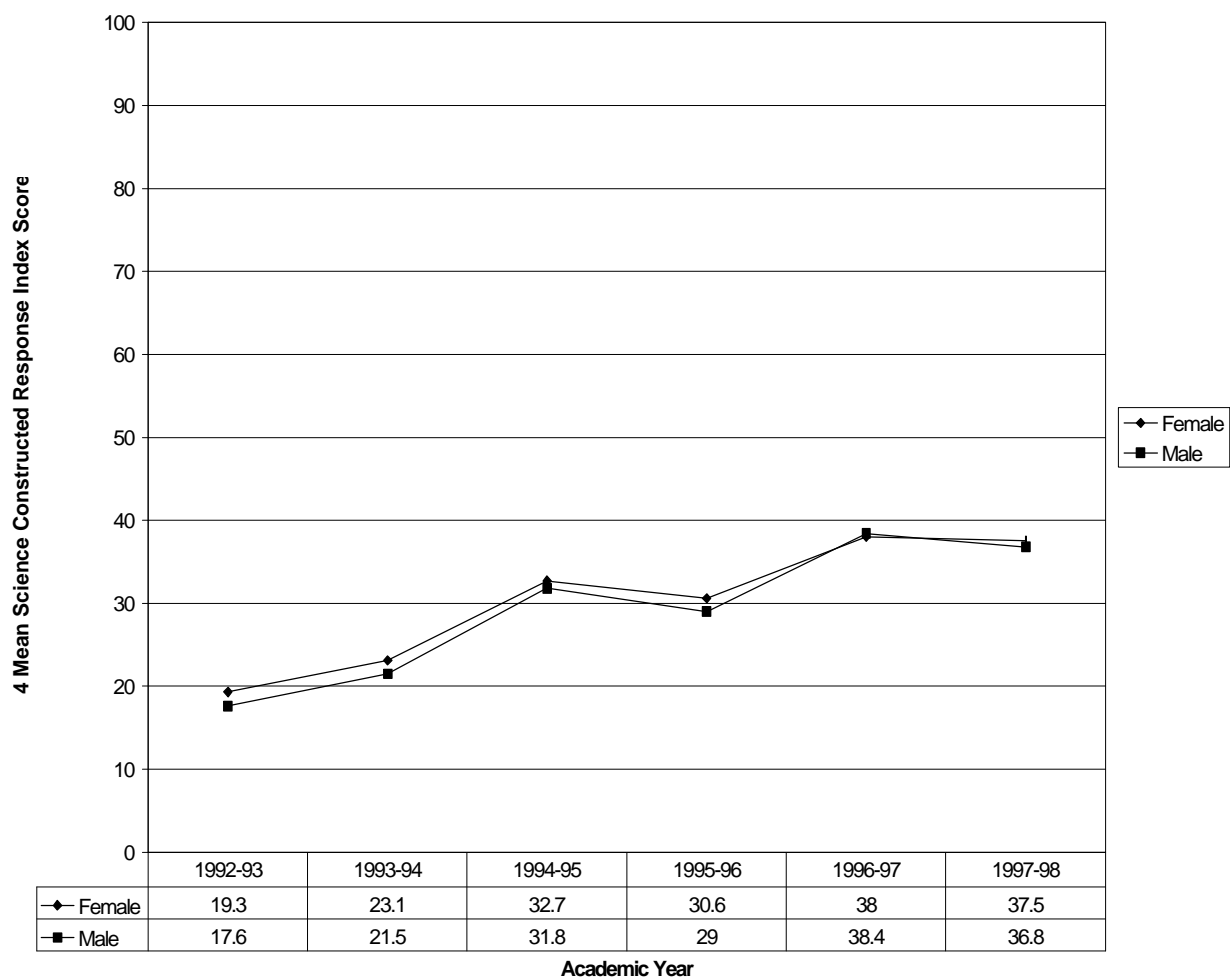


Figure D4. KIRIS Social Studies Constructed Response Index Scores for Grades 4/5 by Gender, 1993-1998

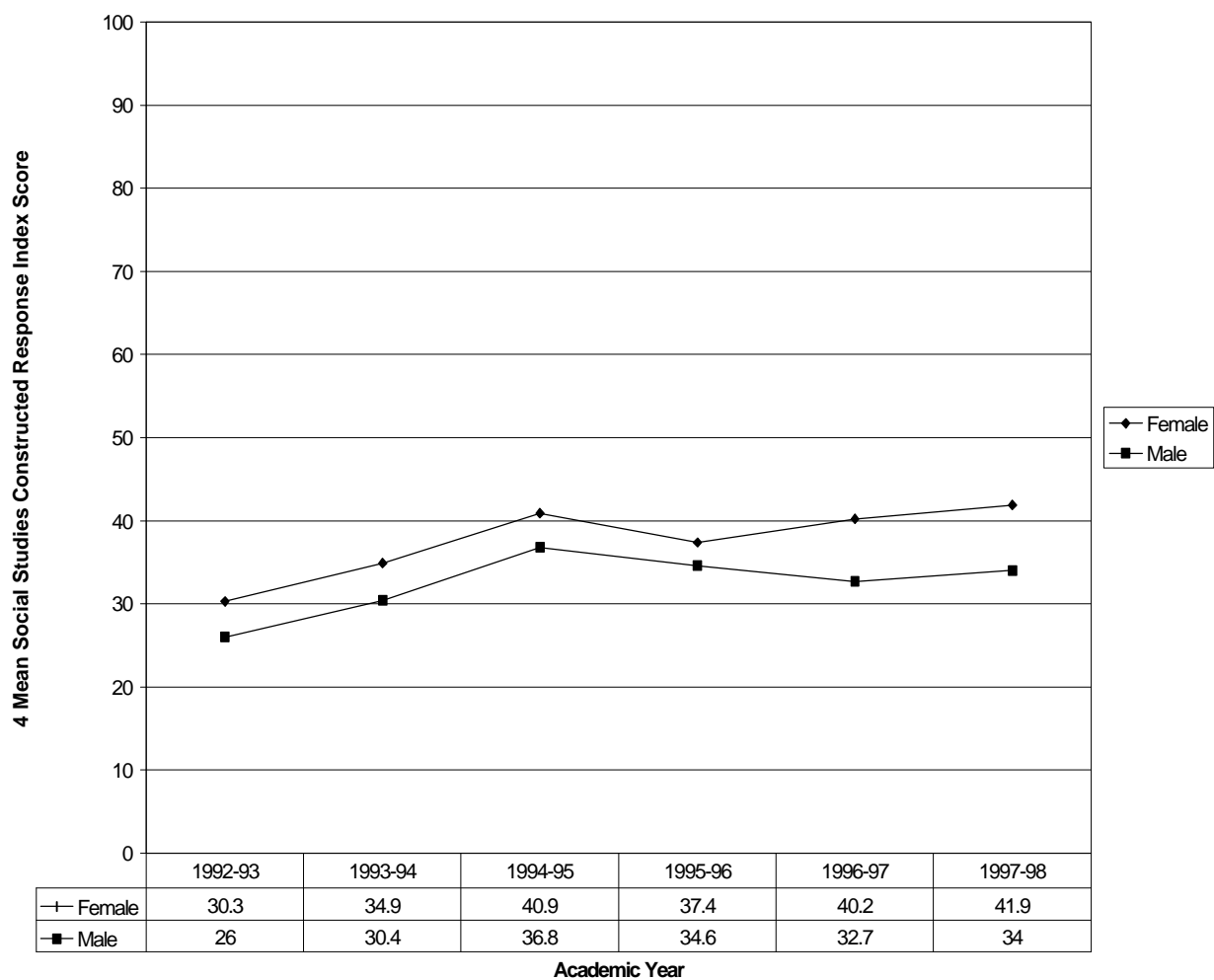


Figure D5. KIRIS Reading Constructed Response Index Scores for Grade 7/8 by Gender, 1993-1998

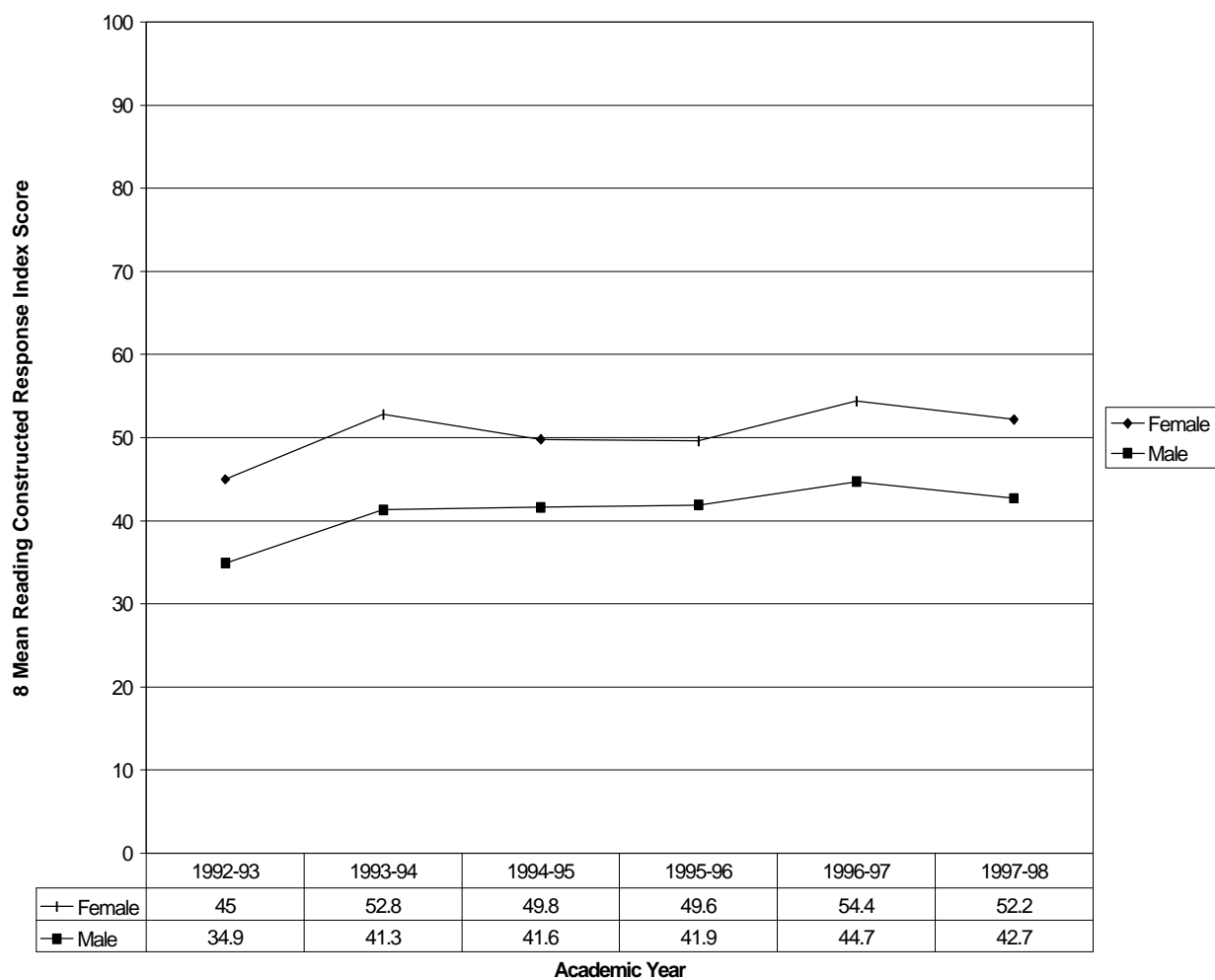


Figure D6. KIRIS Mathematics Constructed Response Index Scores for Grade 8 by Gender, 1993-1998

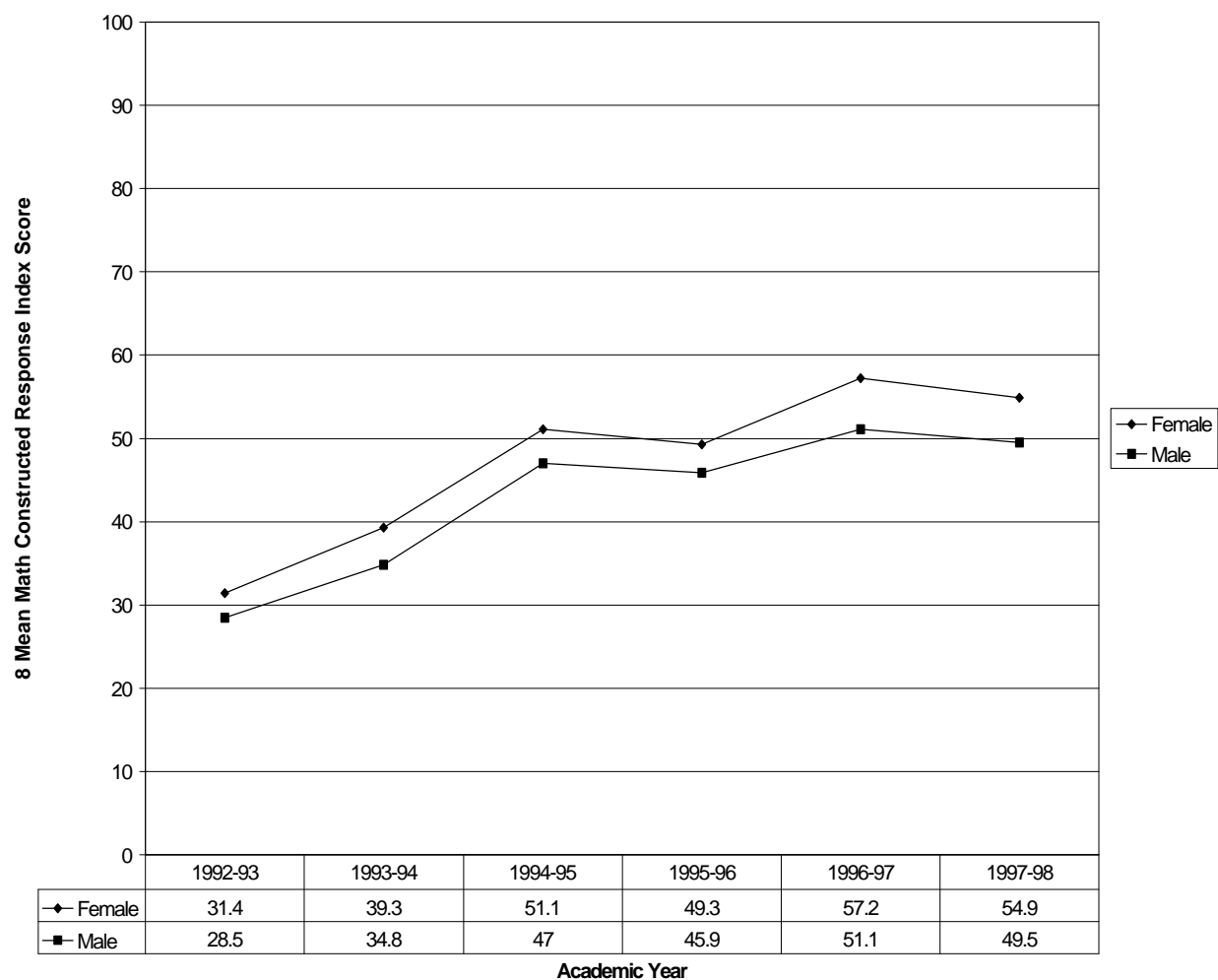


Figure D7. KIRIS Science Constructed Response Index Scores for Grade 7/8 by Gender, 1993-1998

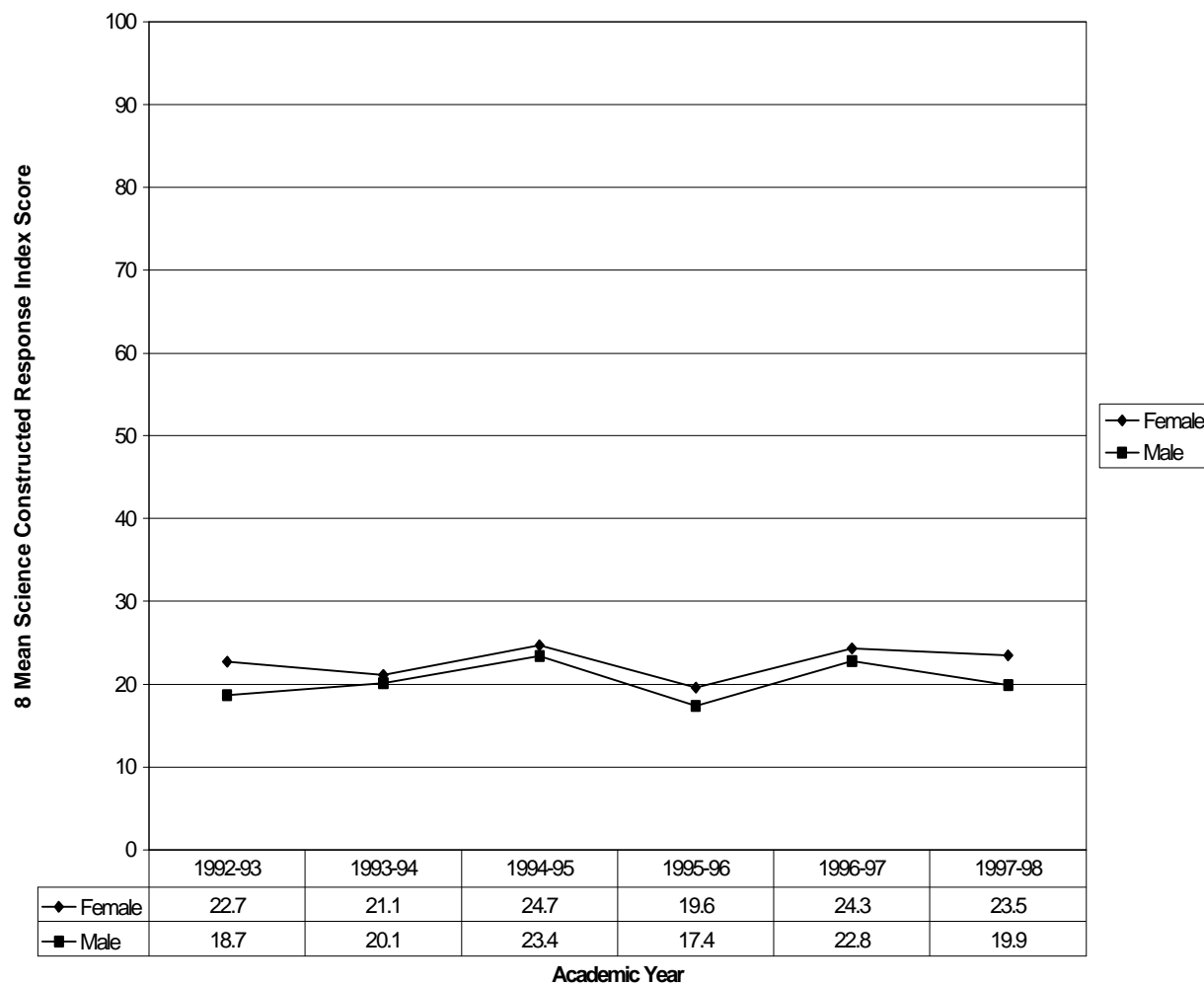


Figure D8. KIRIS Social Studies Constructed Response Index Scores for Grade 8 by Gender, 1993-1998

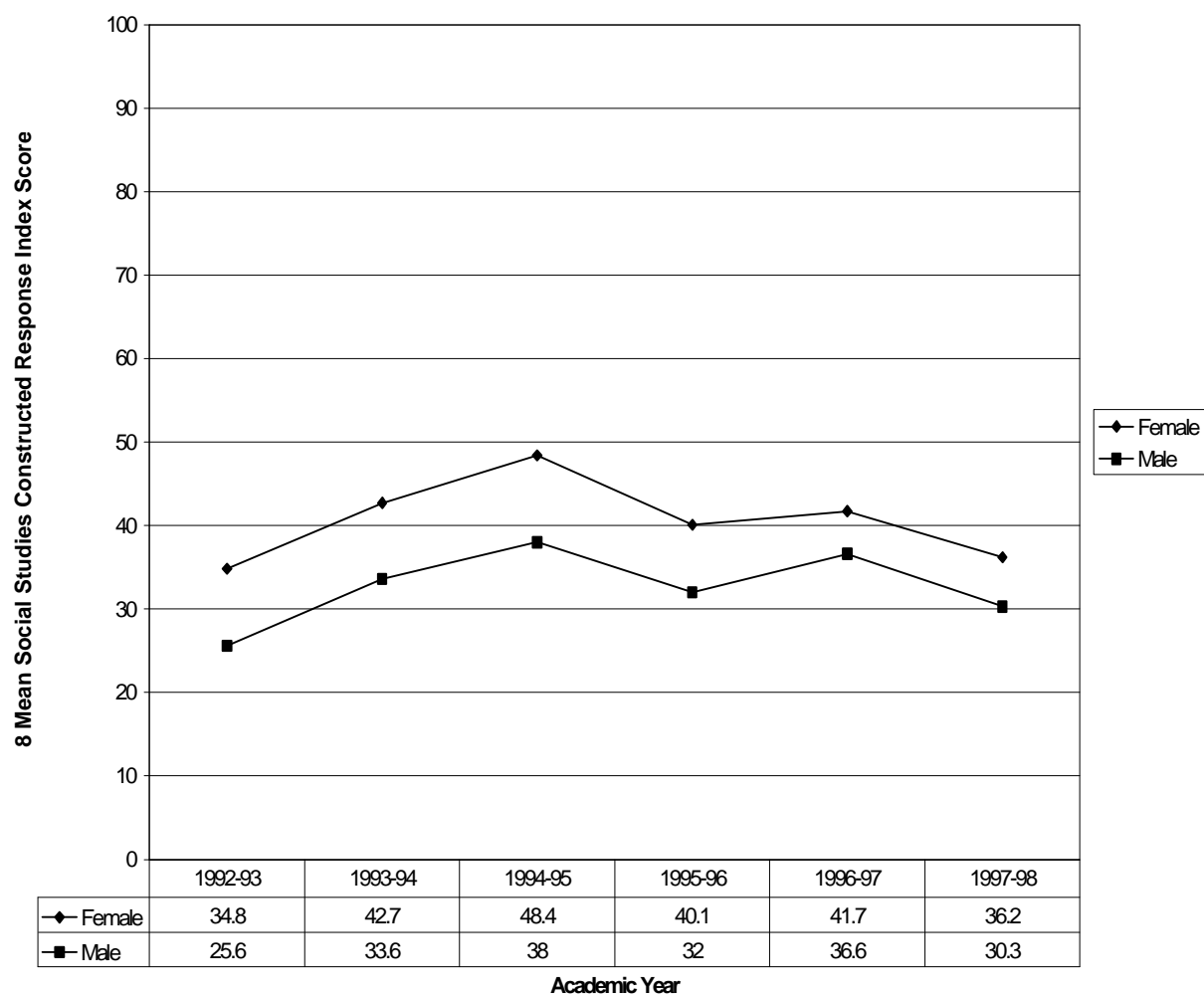


Figure D9. KIRIS Reading Constructed Response Index Scores for Grade 11/12 by Gender, 1993-1998

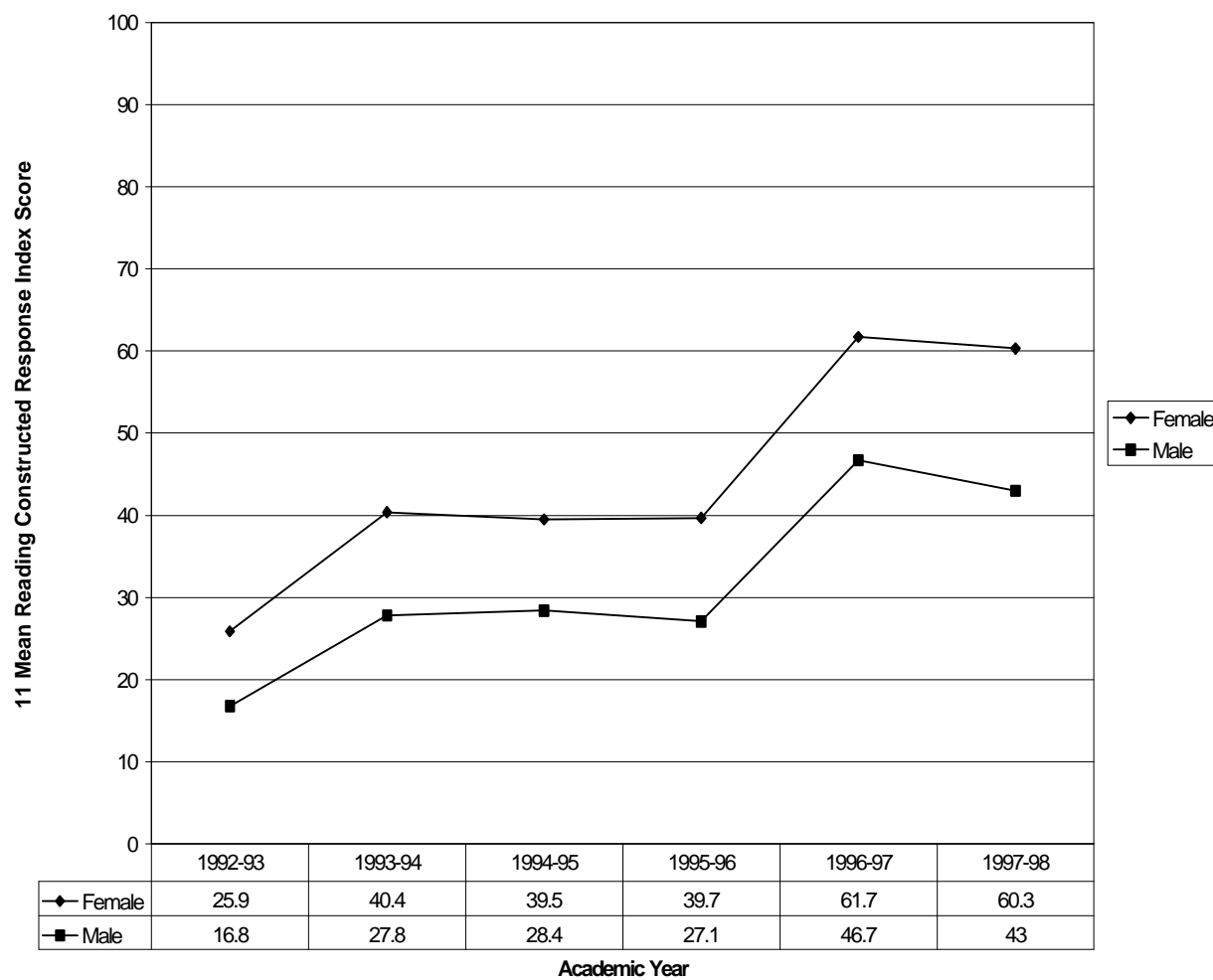


Figure D10. KIRIS Mathematics Constructed Response Index Scores for Grade 11/12 by Gender, 1993-1998

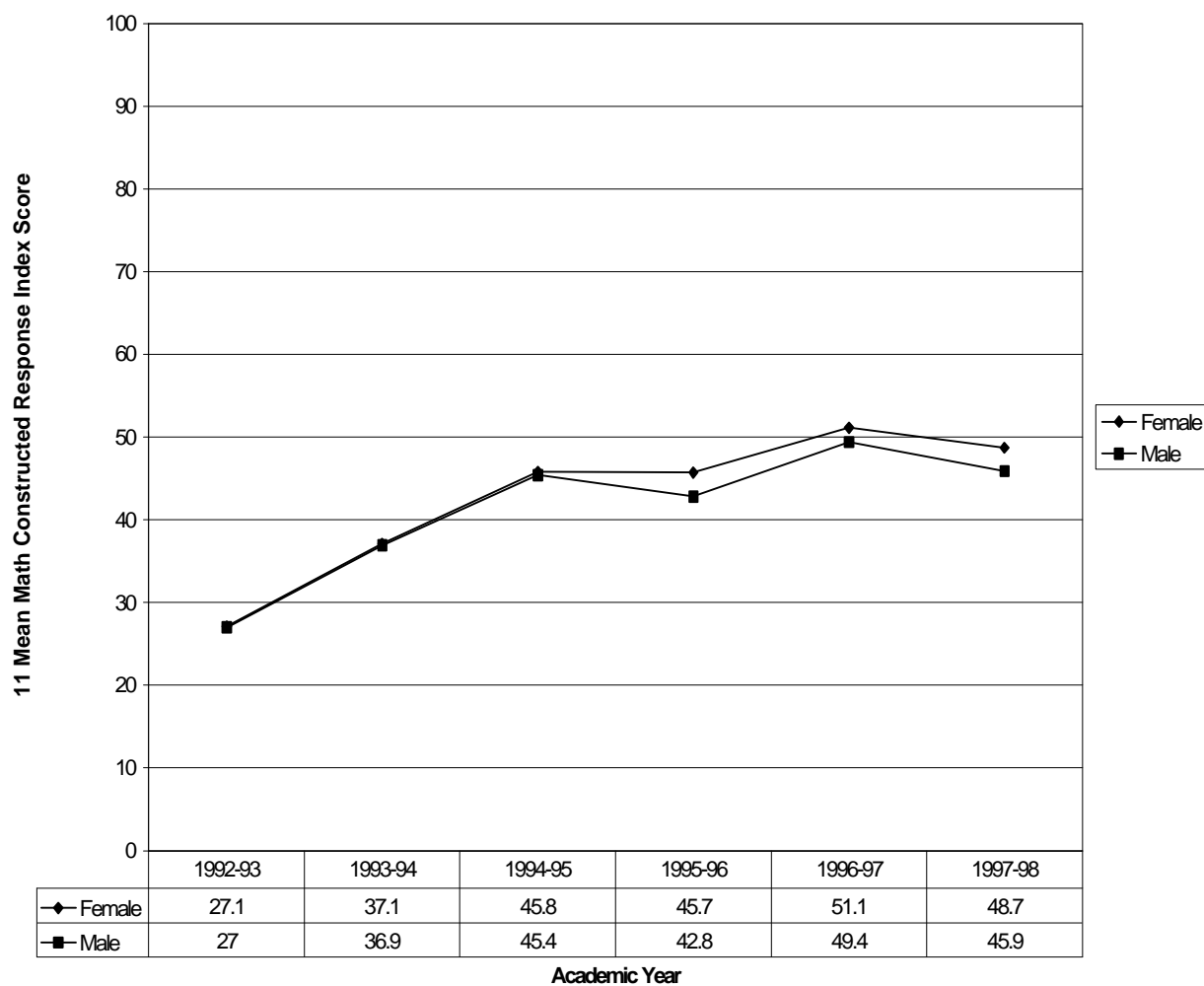


Figure D11. KIRIS Science Constructed Response Theta for Grade 11/12 by Gender, 1993-1998

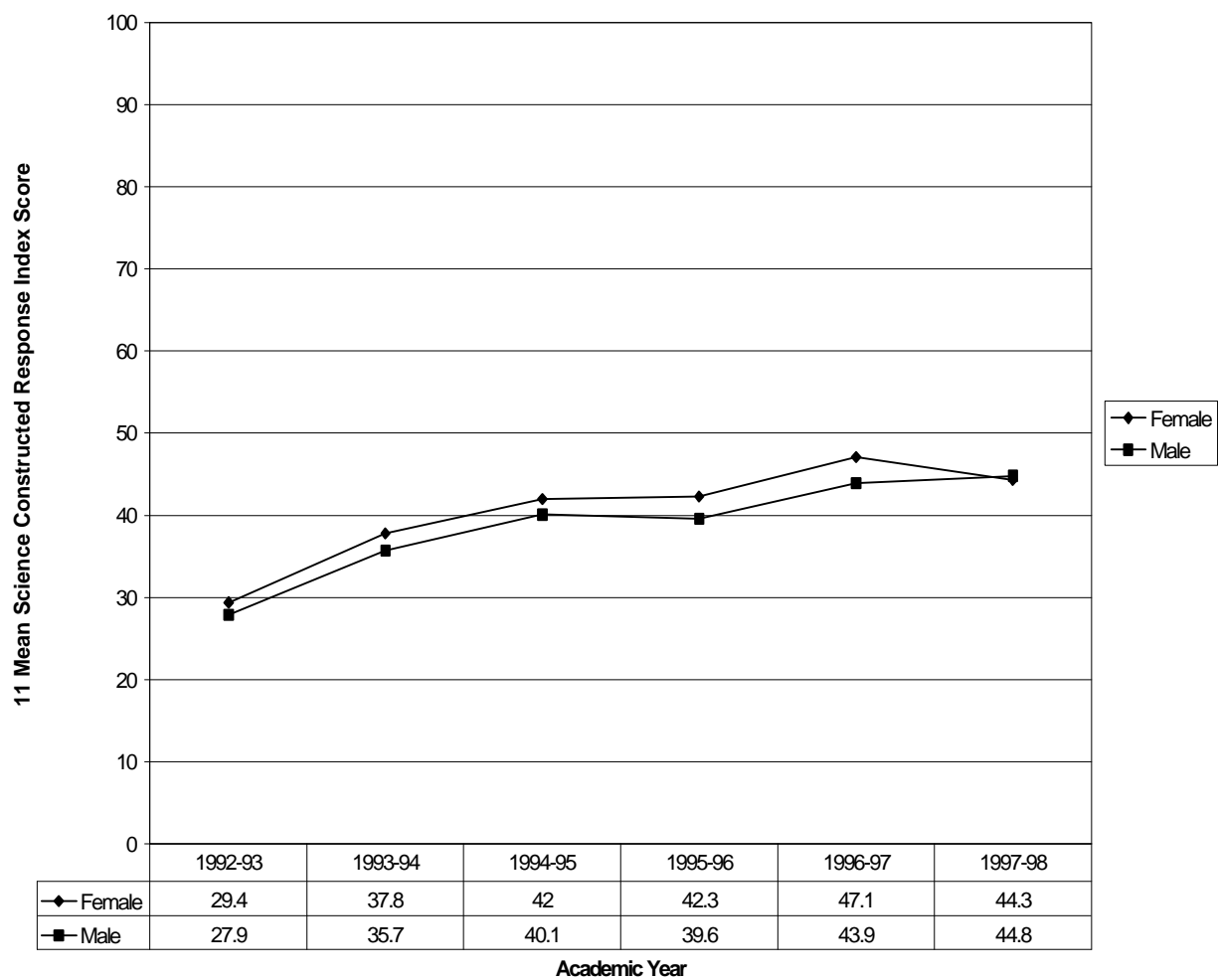
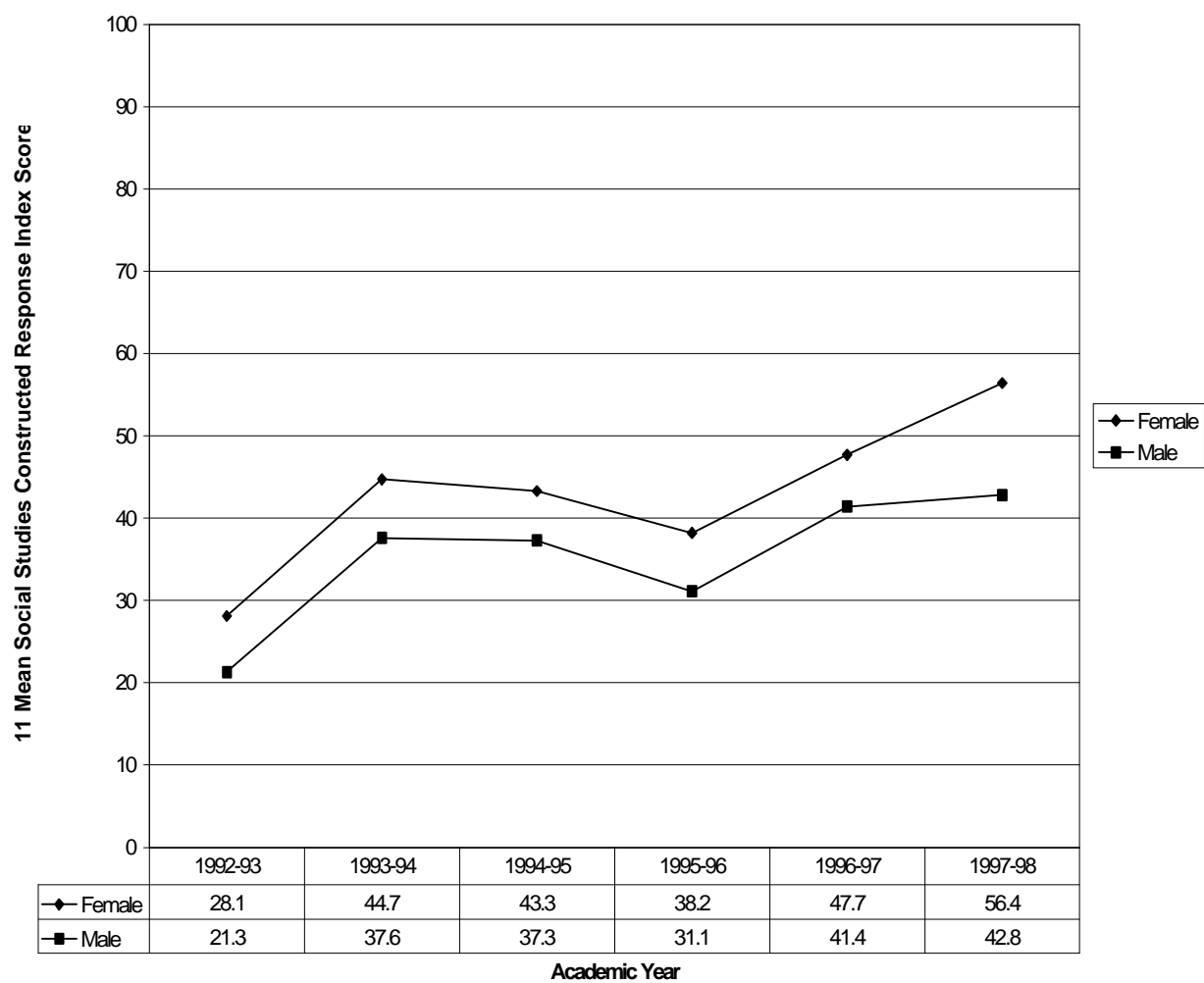


Figure D12. KIRIS Social Studies Constructed Response Index Scores for Grade 11/12 by Gender, 1993-1998



Appendix E⁴⁵

KIRIS Cycle 2 and Cycle 3 Subject Area Performance as
Measured by Constructed Response Index Scores for All
Grades Tested by Race

⁴⁵ Data Tables were added to all Figures in Appendix E by KDE.

Figure E1. KIRIS Reading Constructed Response Index Scores for Grade 4 by Race, 1993-1998

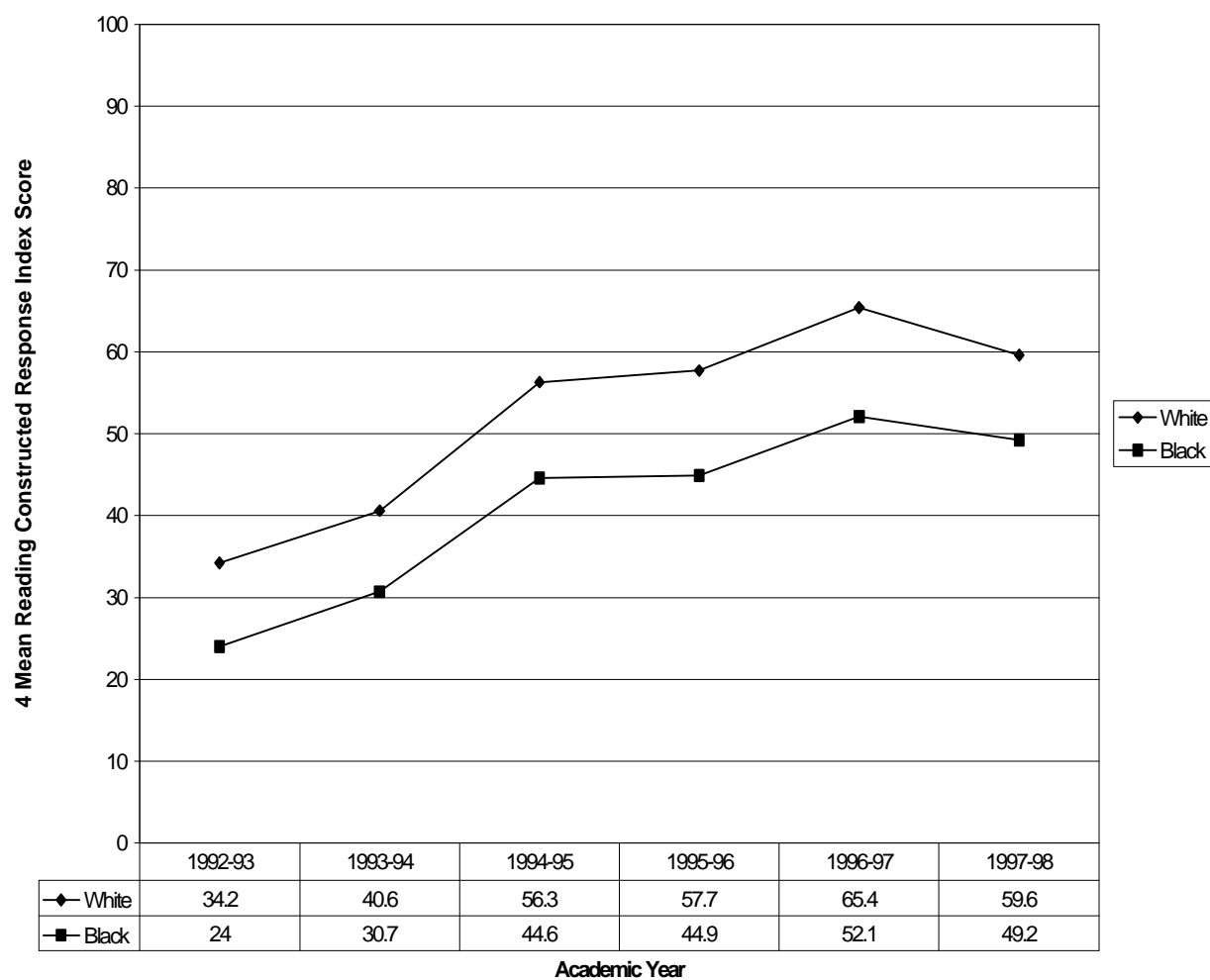


Figure E2. KIRIS Mathematics Constructed Response Index Scores for Grade 4/5 by Race, 1993-1998

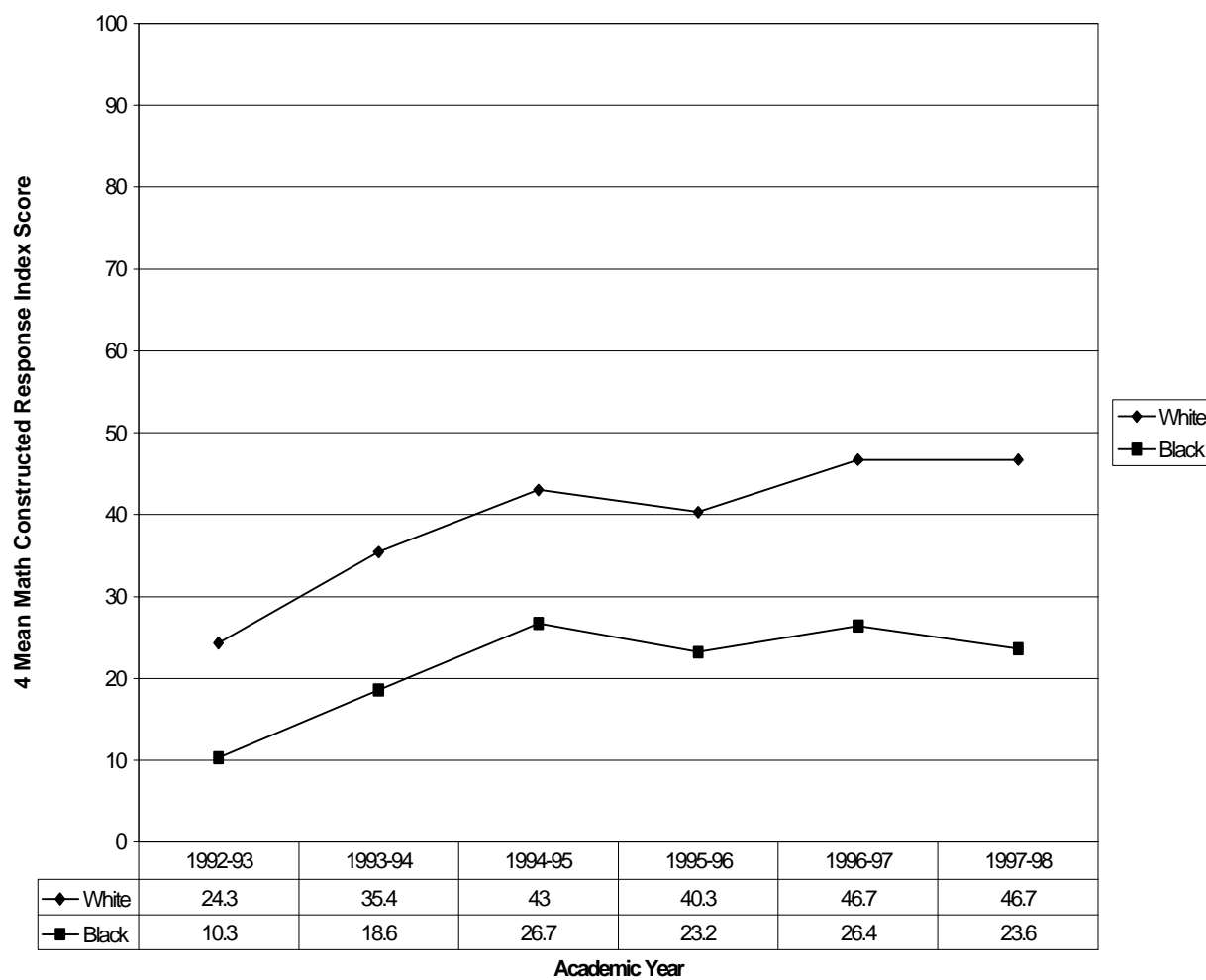


Figure E3. KIRIS Science Constructed Response Index Scores for Grade 4 by Race, 1993-1998

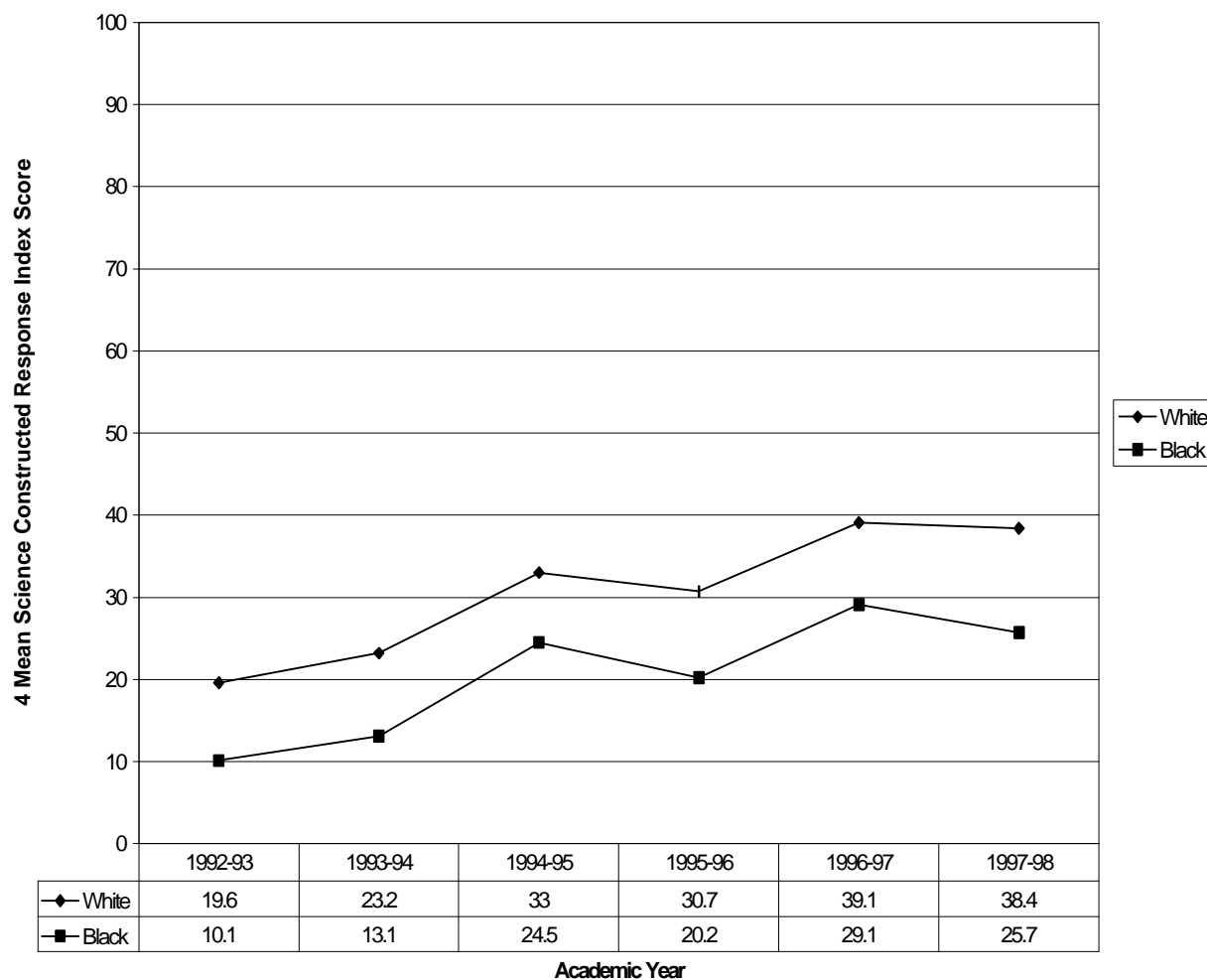


Figure E4. KIRIS Social Studies Constructed Response Index Scores for Grade 4/5 by Race, 1993-1998

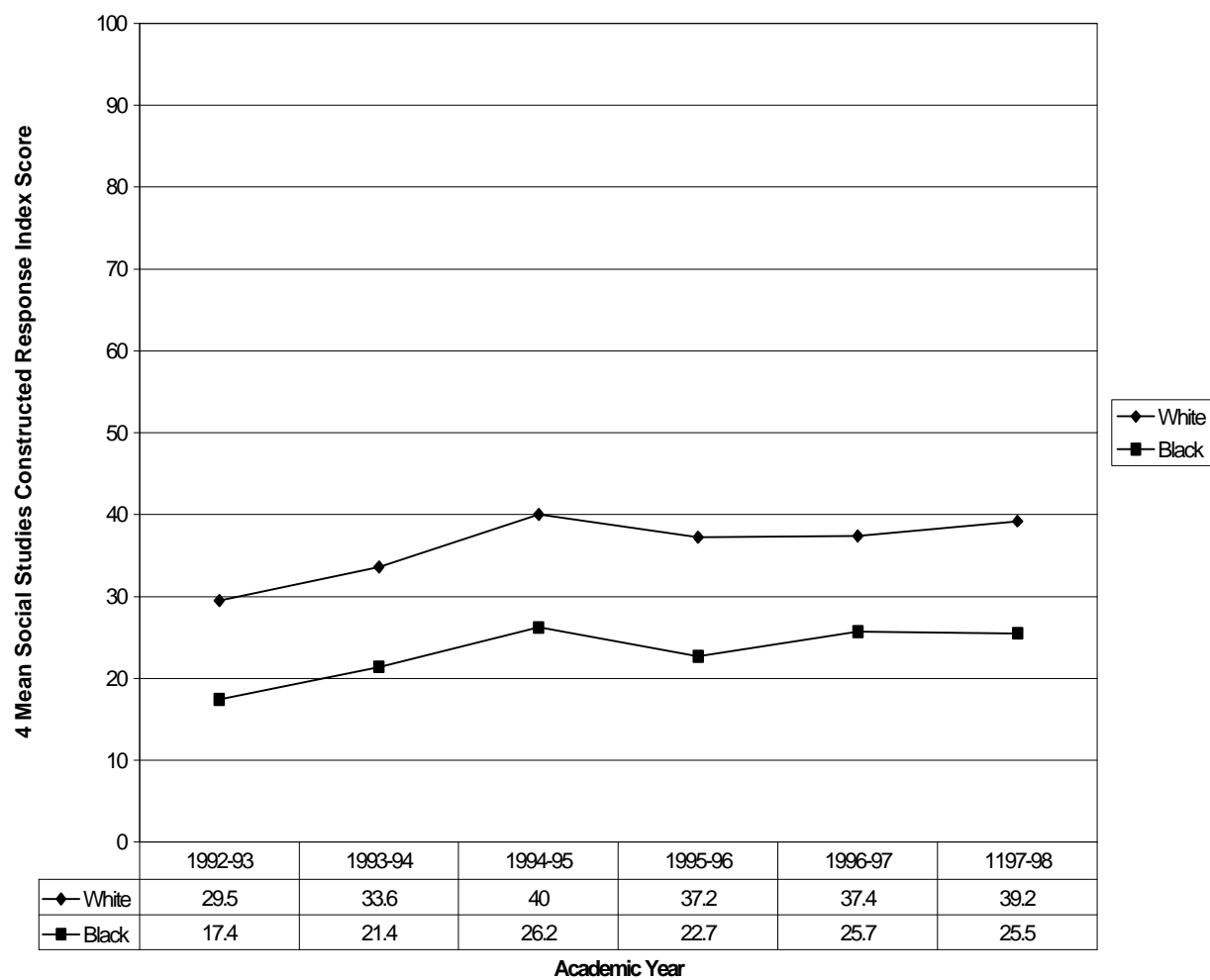


Figure E5. KIRIS Reading Constructed Response Index Scores for Grades 7/8 by Race, 1993-1998

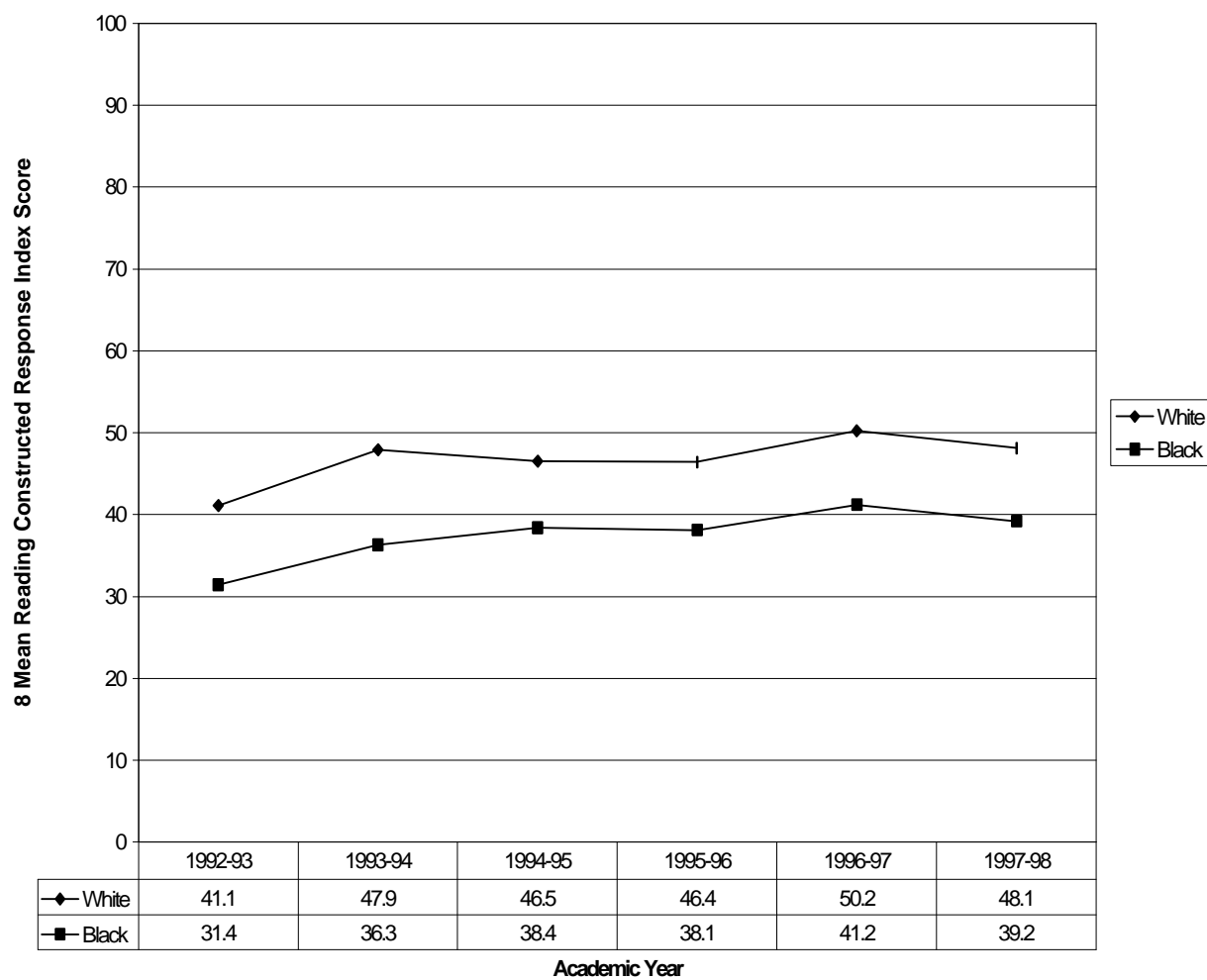


Figure E6. KIRIS Mathematics Constructed Response Index Scores for Grade 8 by Race, 1993-1998

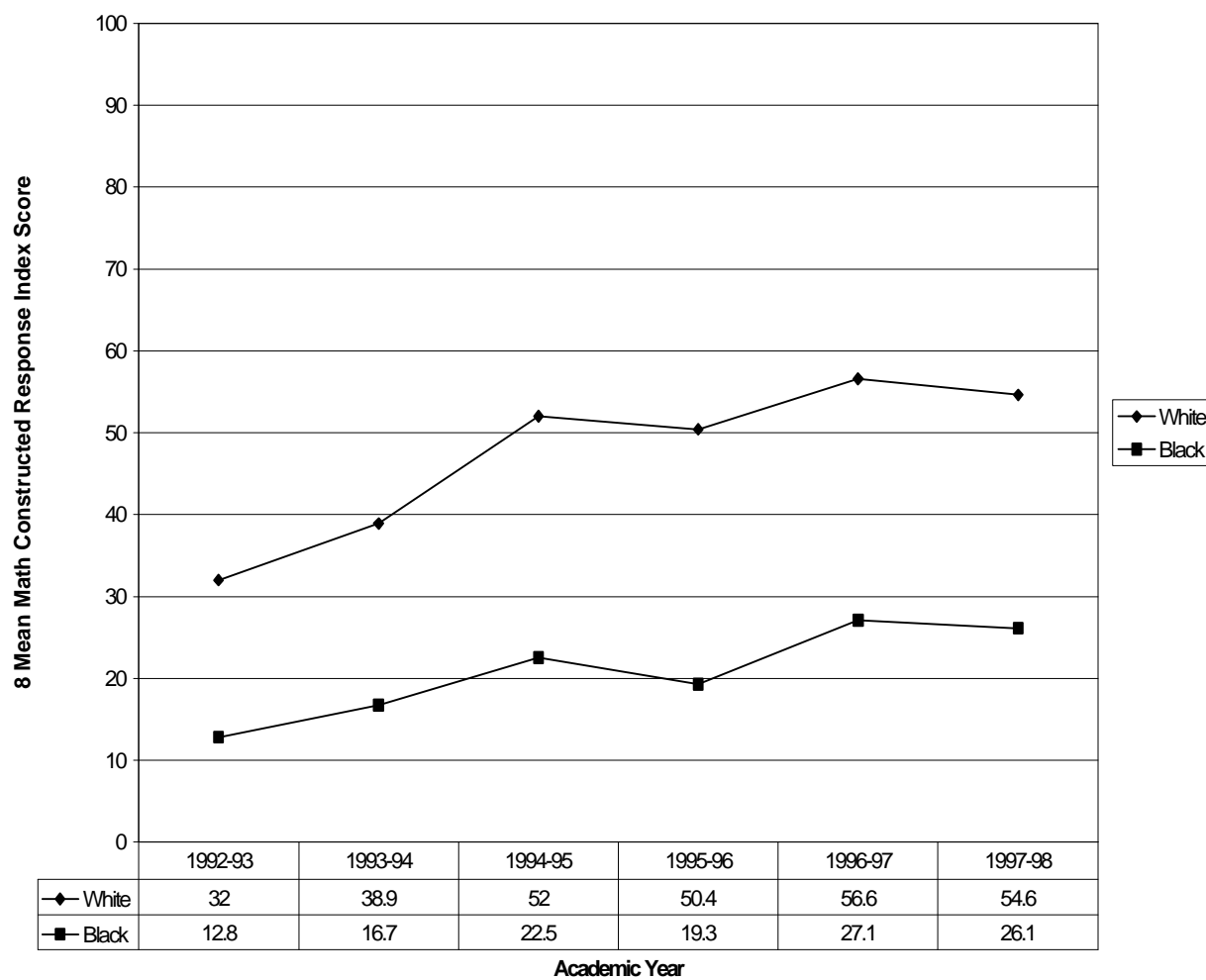


Figure E7. KIRIS Science Constructed Response Index Scores for Grade 7/8 by Race, 1993-1998

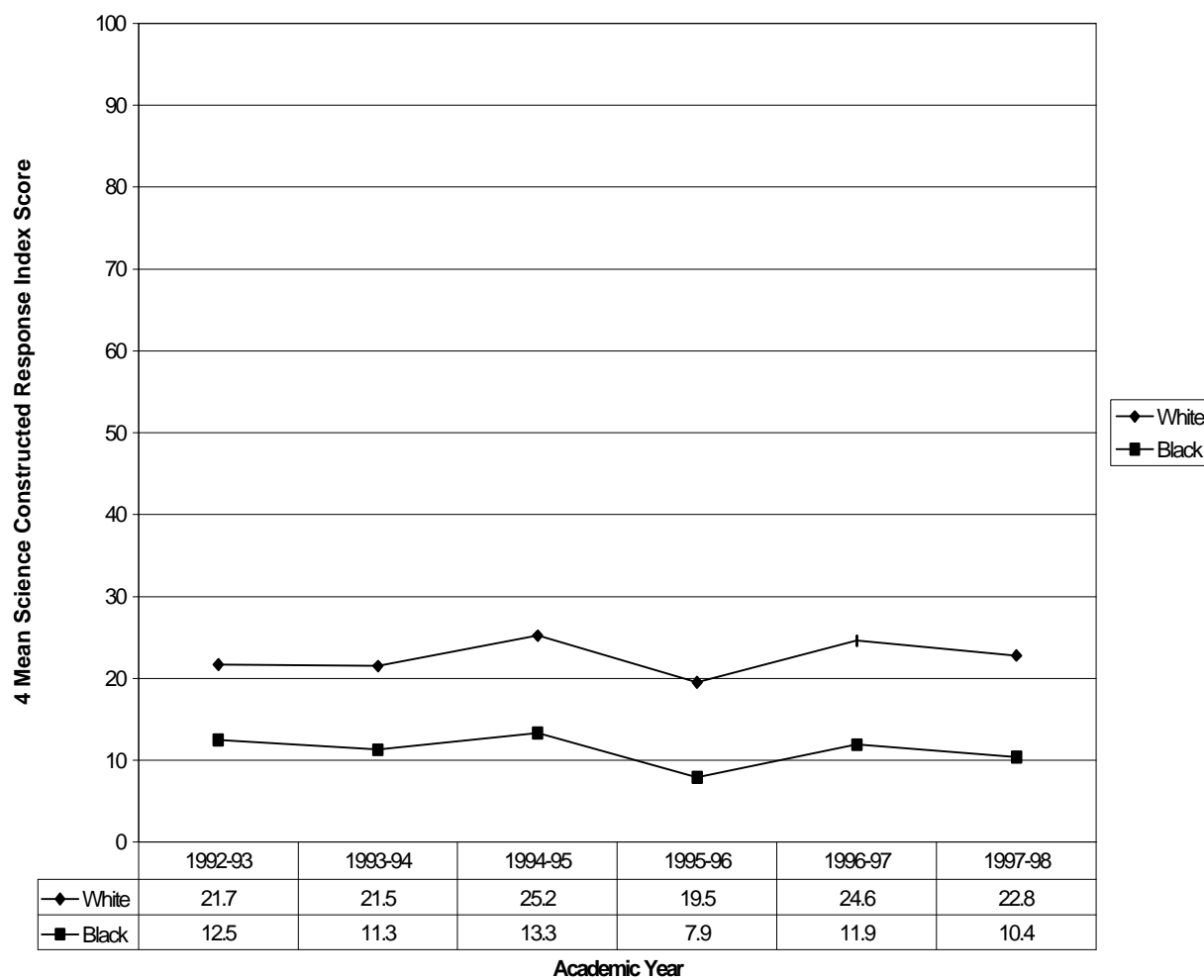


Figure E8. KIRIS Social Studies Constructed Response Index for Grade 8 by Race, 1993-1998

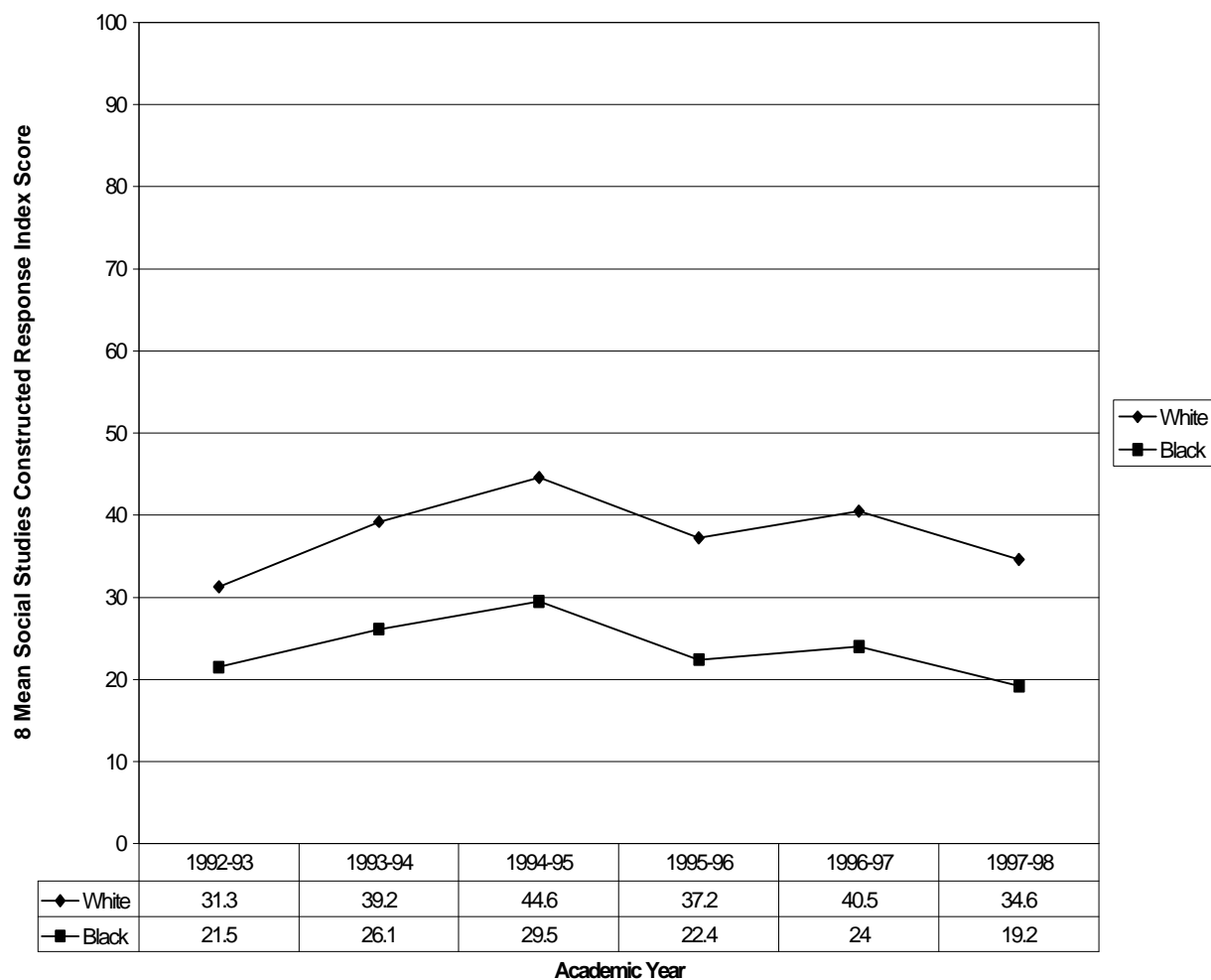


Figure E9. KIRIS Reading Constructed Response Index Scores for Grade 11/12 by Race, 1993-1998

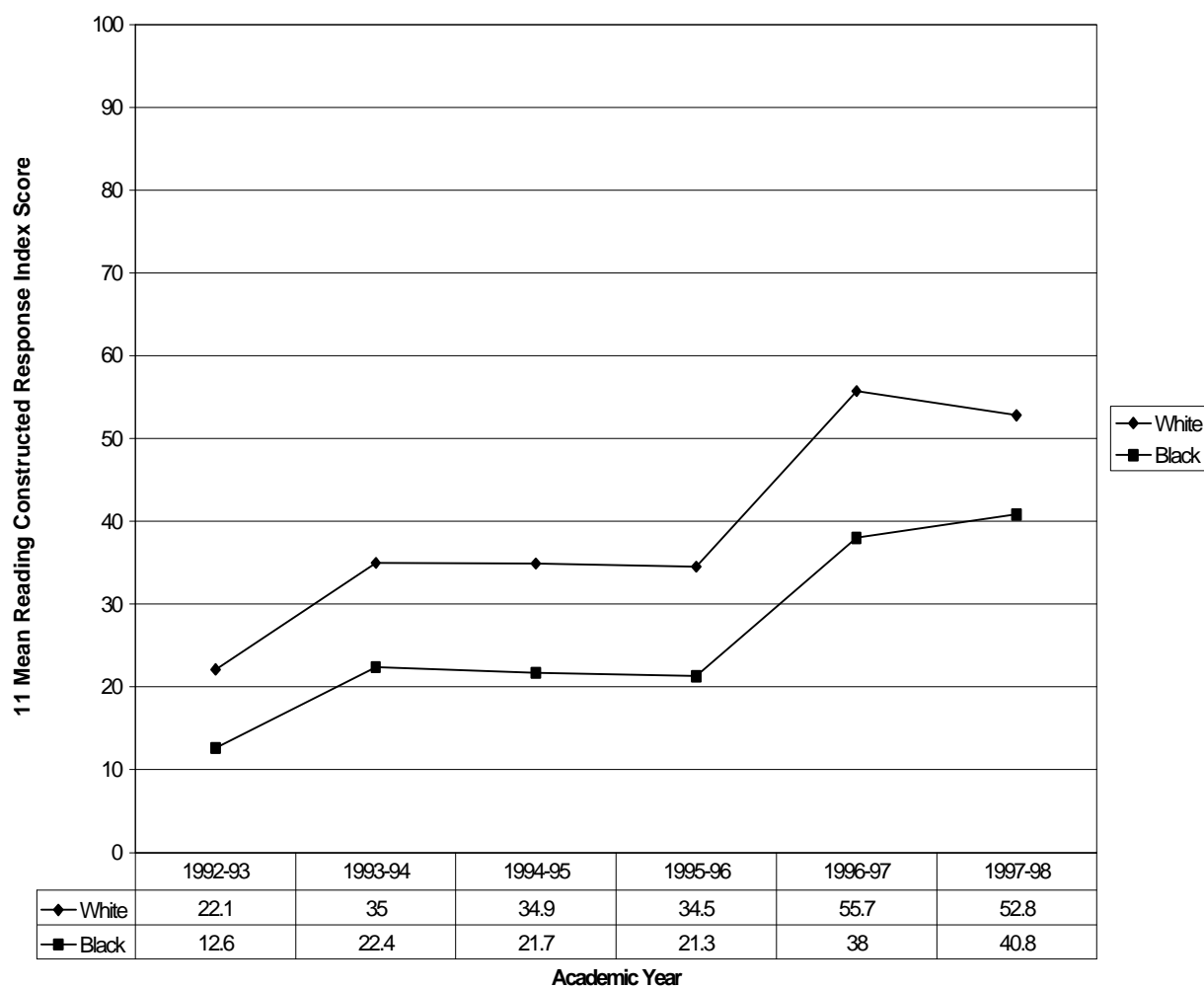


Figure E10. KIRIS Mathematics Constructed Response Index Scores for Grade 11/12 by Race, 1993-1998

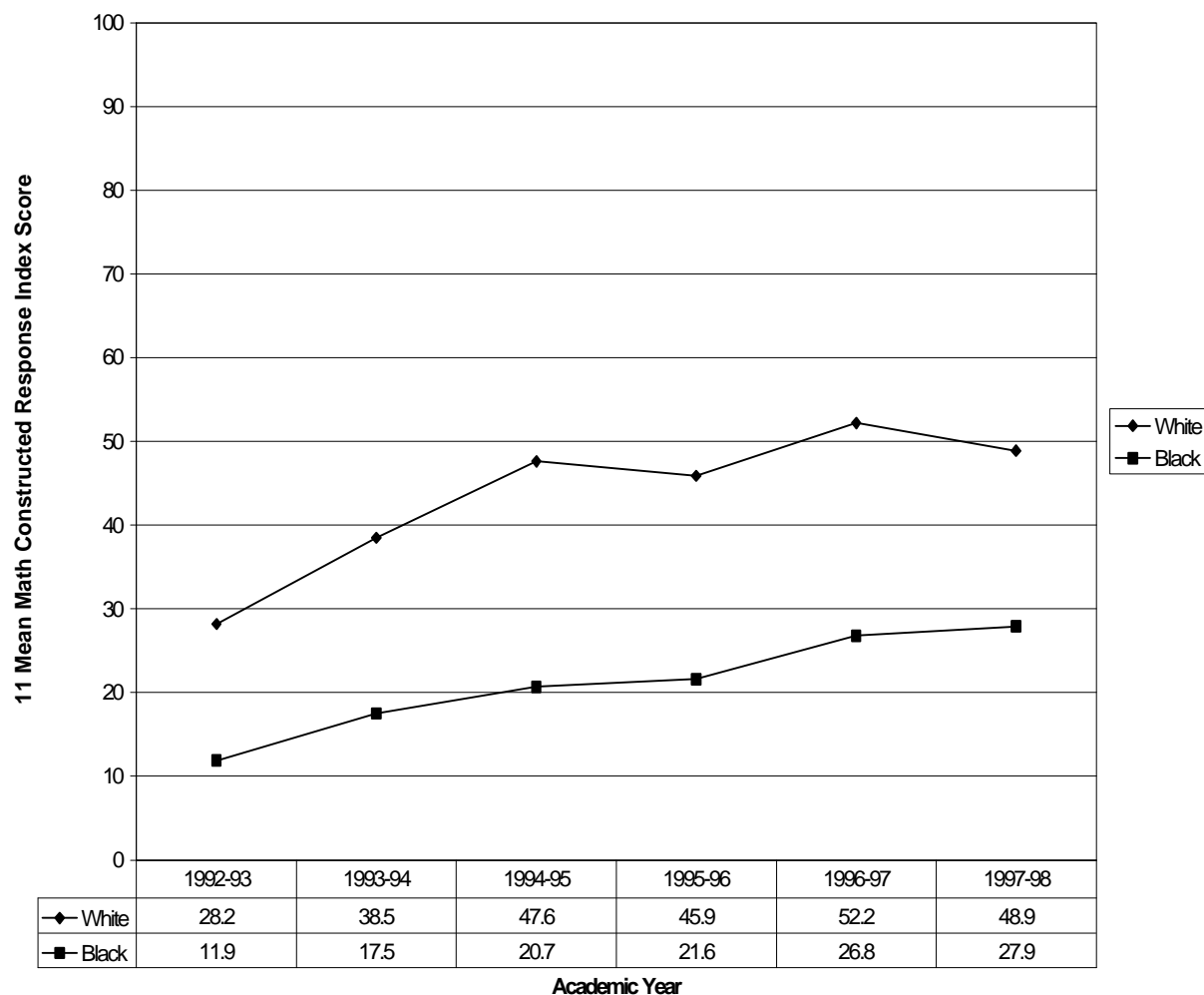


Figure E11. KIRIS Science Constructed Response Index Scores for Grade 11/12 by Race, 1993-1998

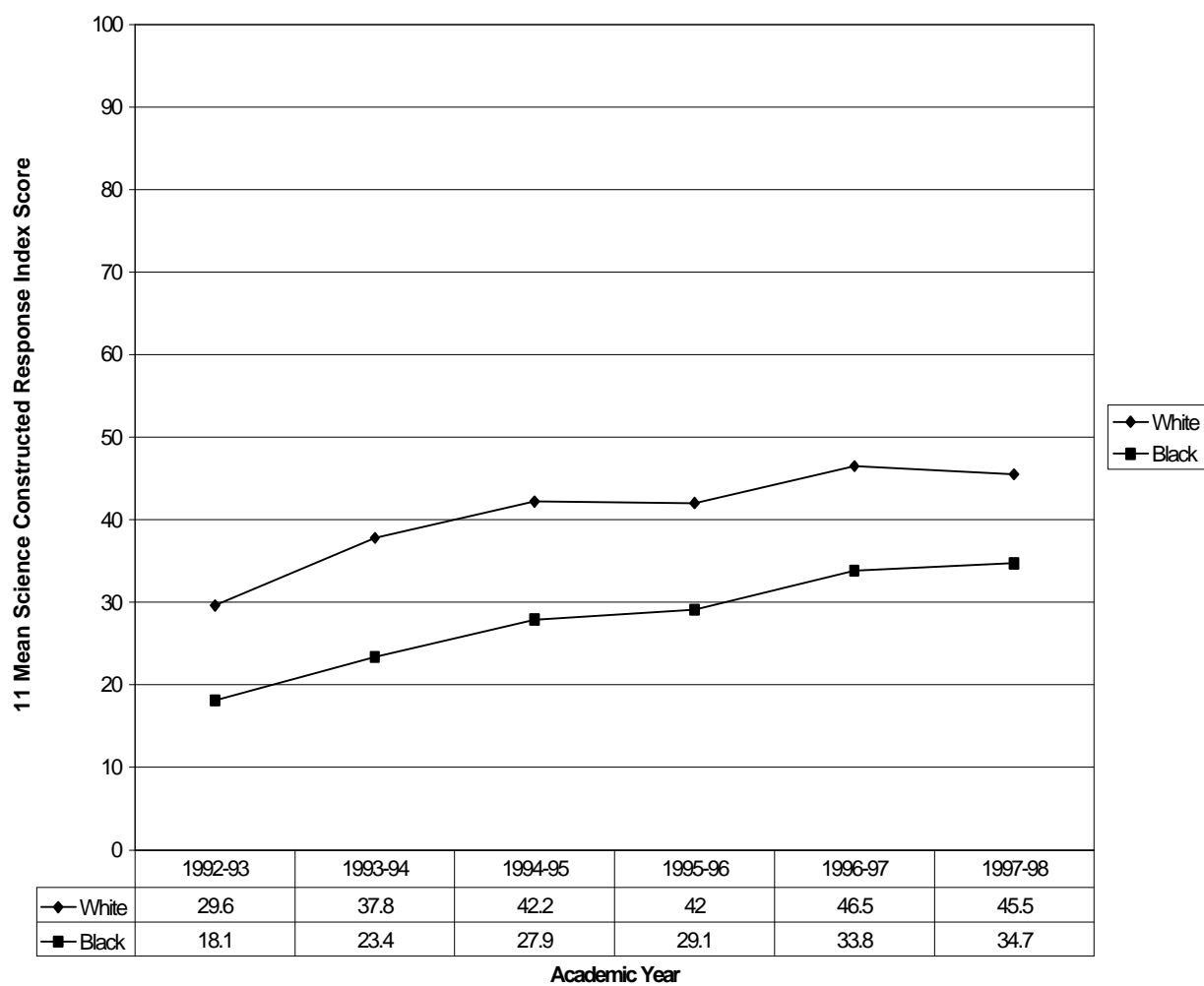
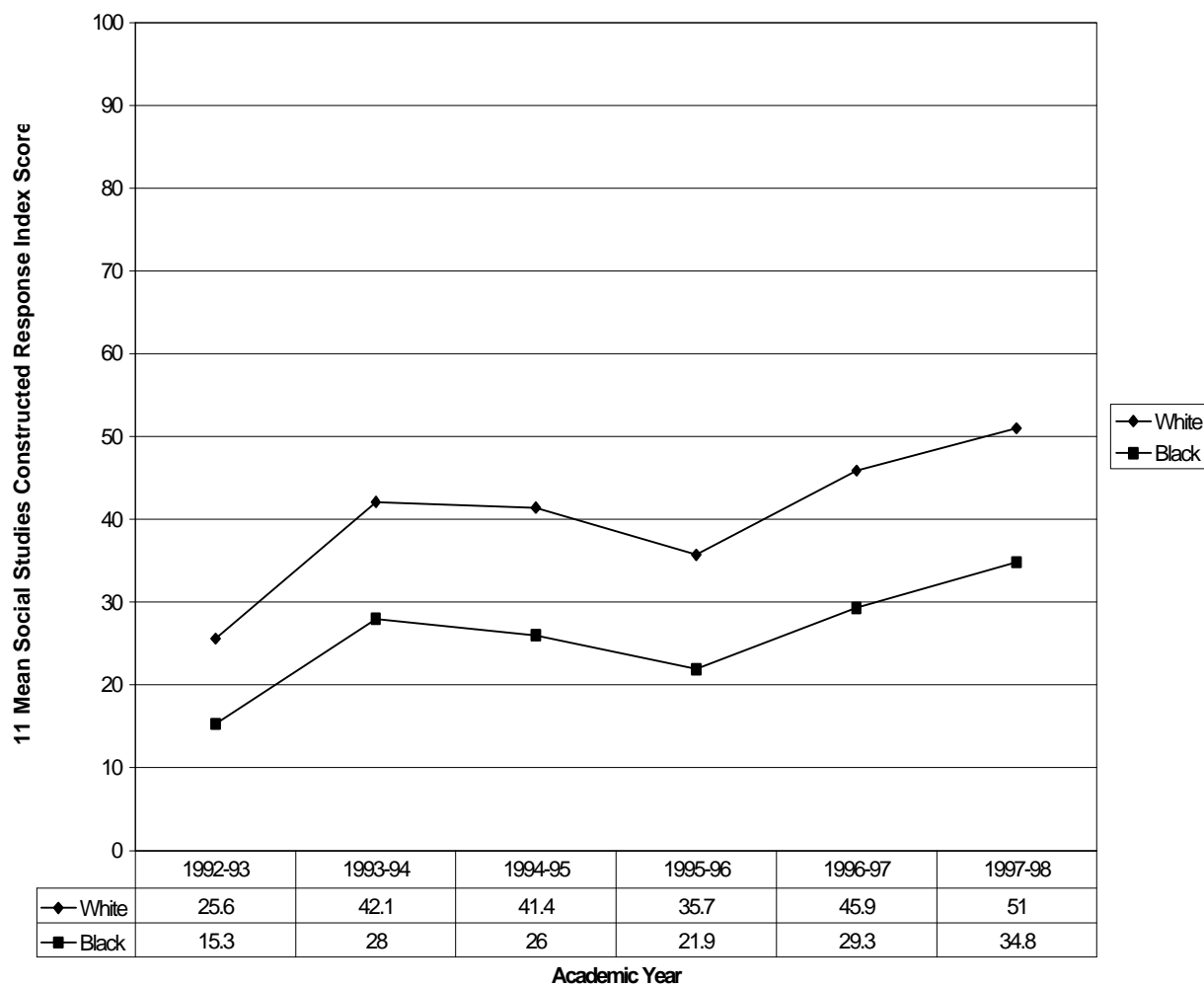


Figure E12. KIRIS Social Studies Constructed Response Index Scores for Grade 11/12 by Race, 1993-1998



Appendix F⁴⁶

KIRIS Cycle 2 and Cycle 3 Subject Area Performance for Males
by Constructed Response Index Categories

⁴⁶ Data Tables were added to all Figures in Appendix F by KDE.

Figure F1. Males' Percentage NAPD in Grade 4 Reading, 1993-1998

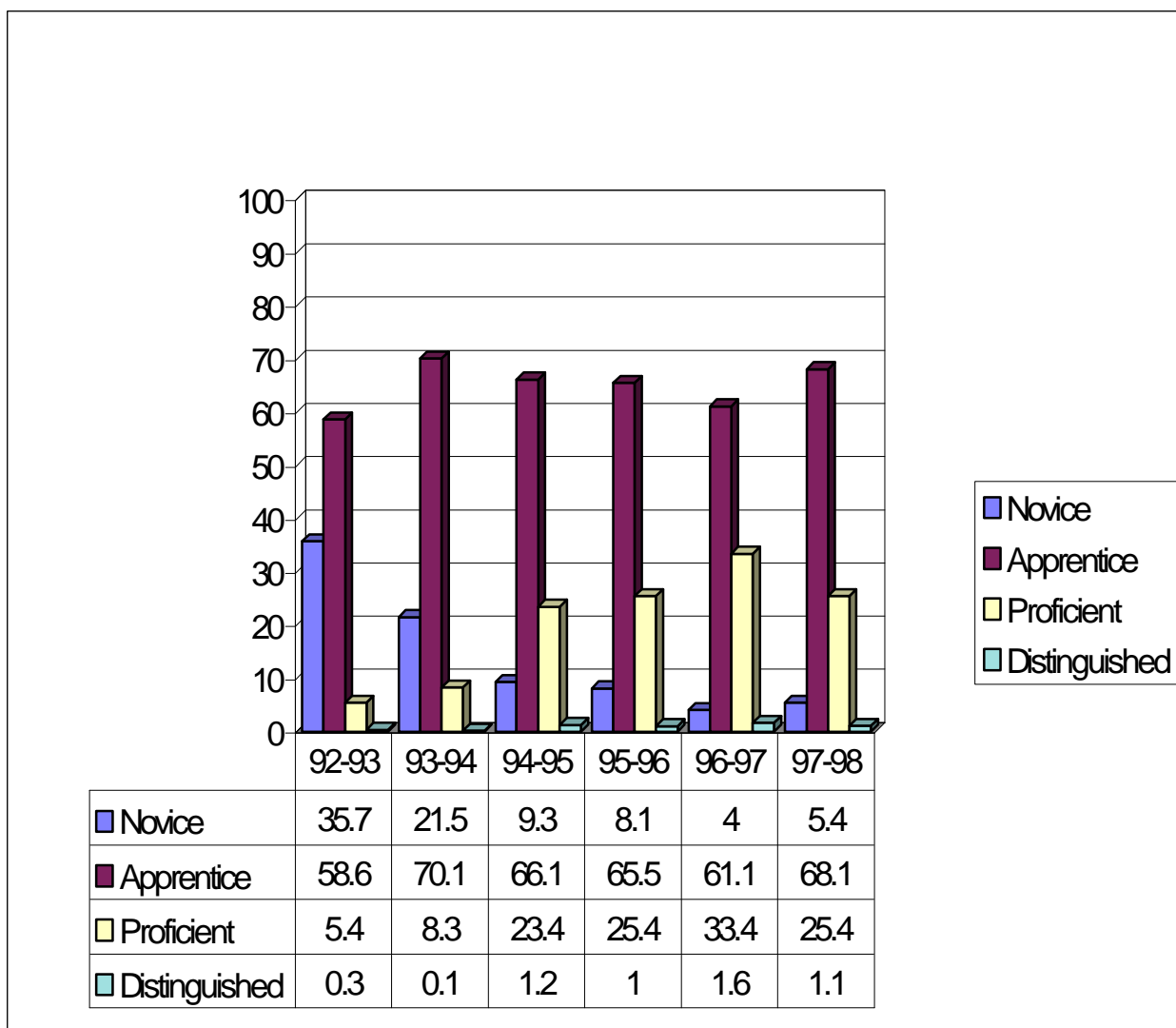


Figure F2 Males' Percentage NAPD in Grade 4/5 Math, 1993-1998

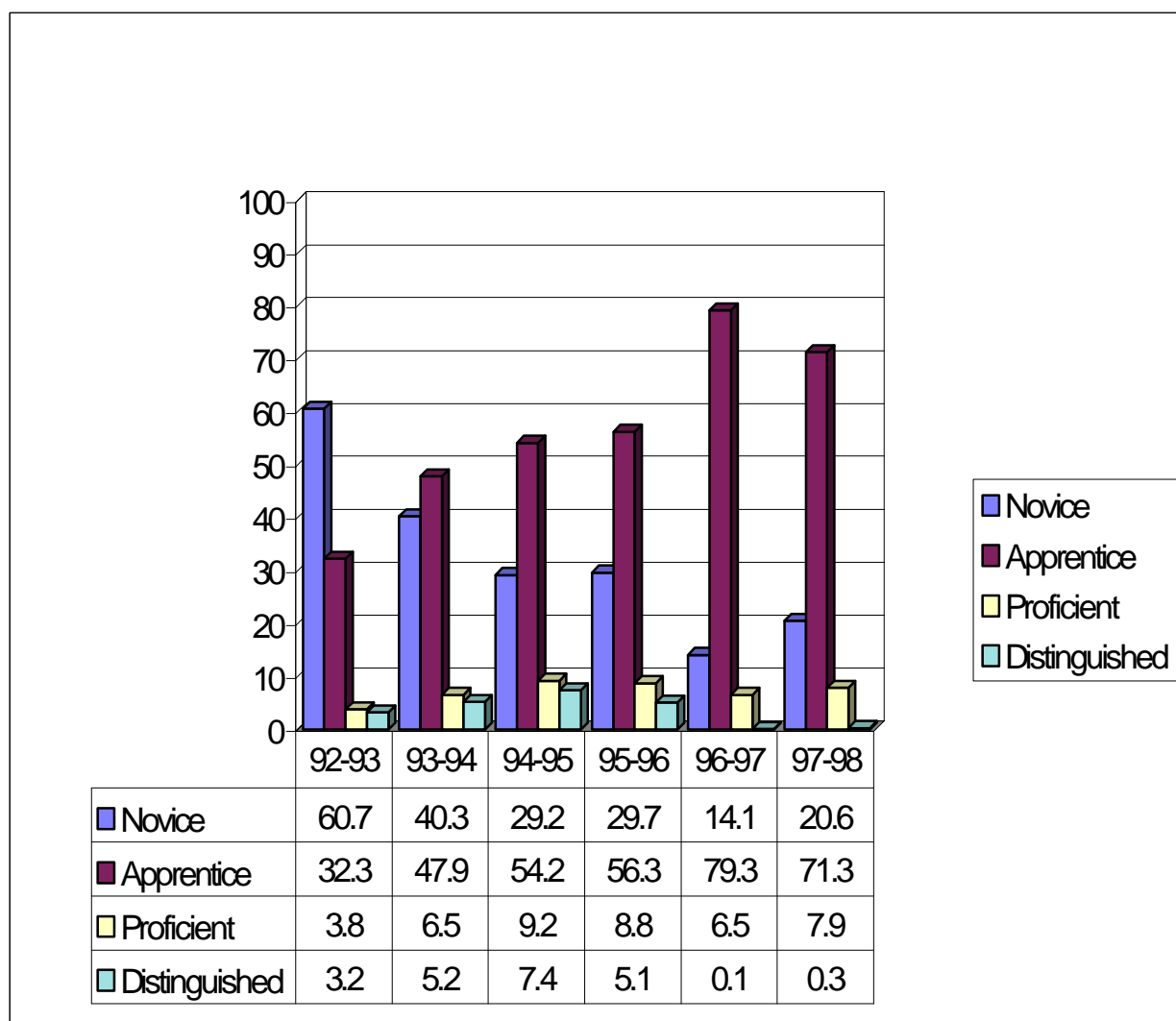


Figure F3. Males' Percentage NAPD in Grade 4 Science, 1993-1998

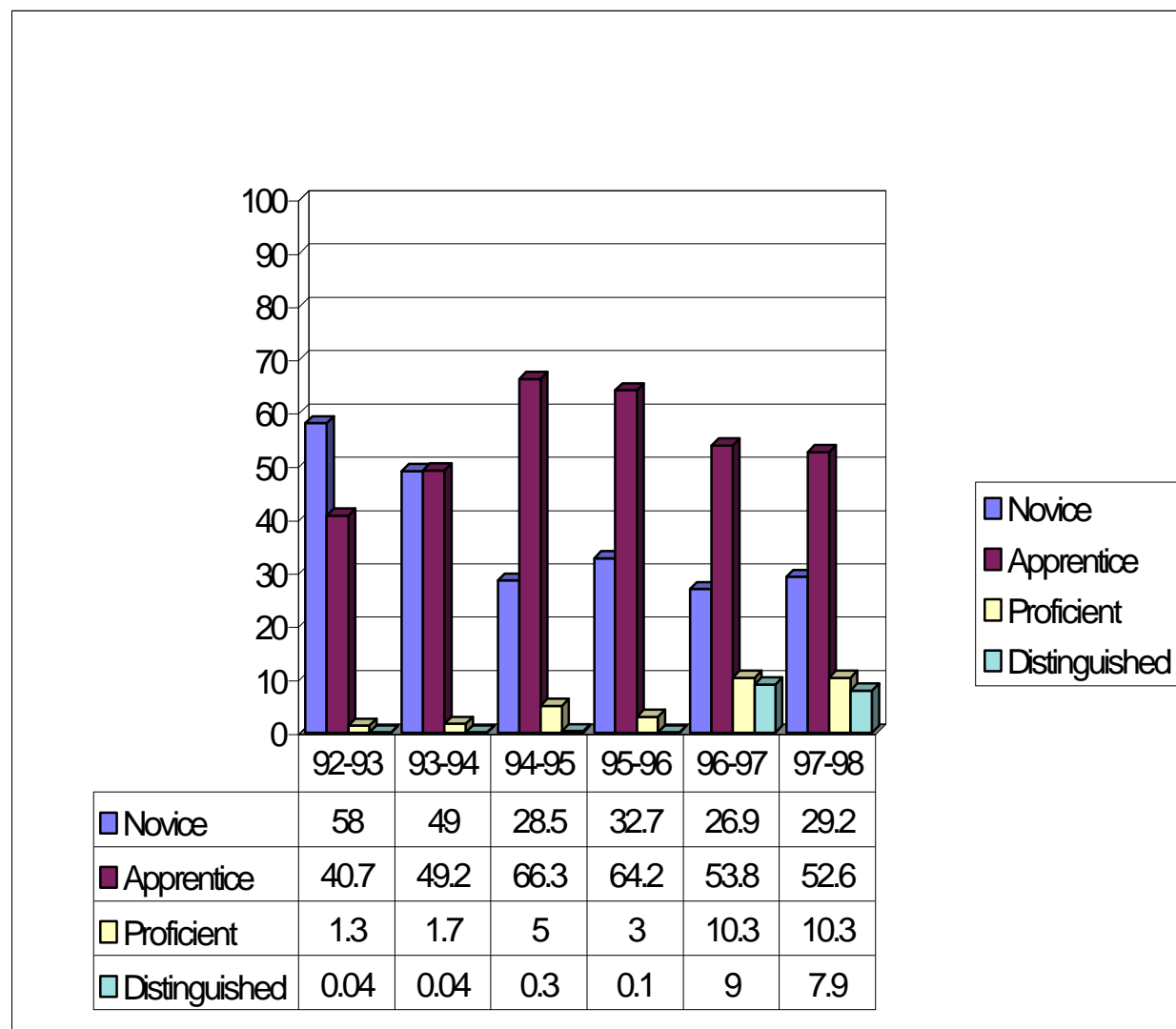


Figure F4. Males' Percentage NAPD in Grade 4/5 Social Studies, 1993-1998

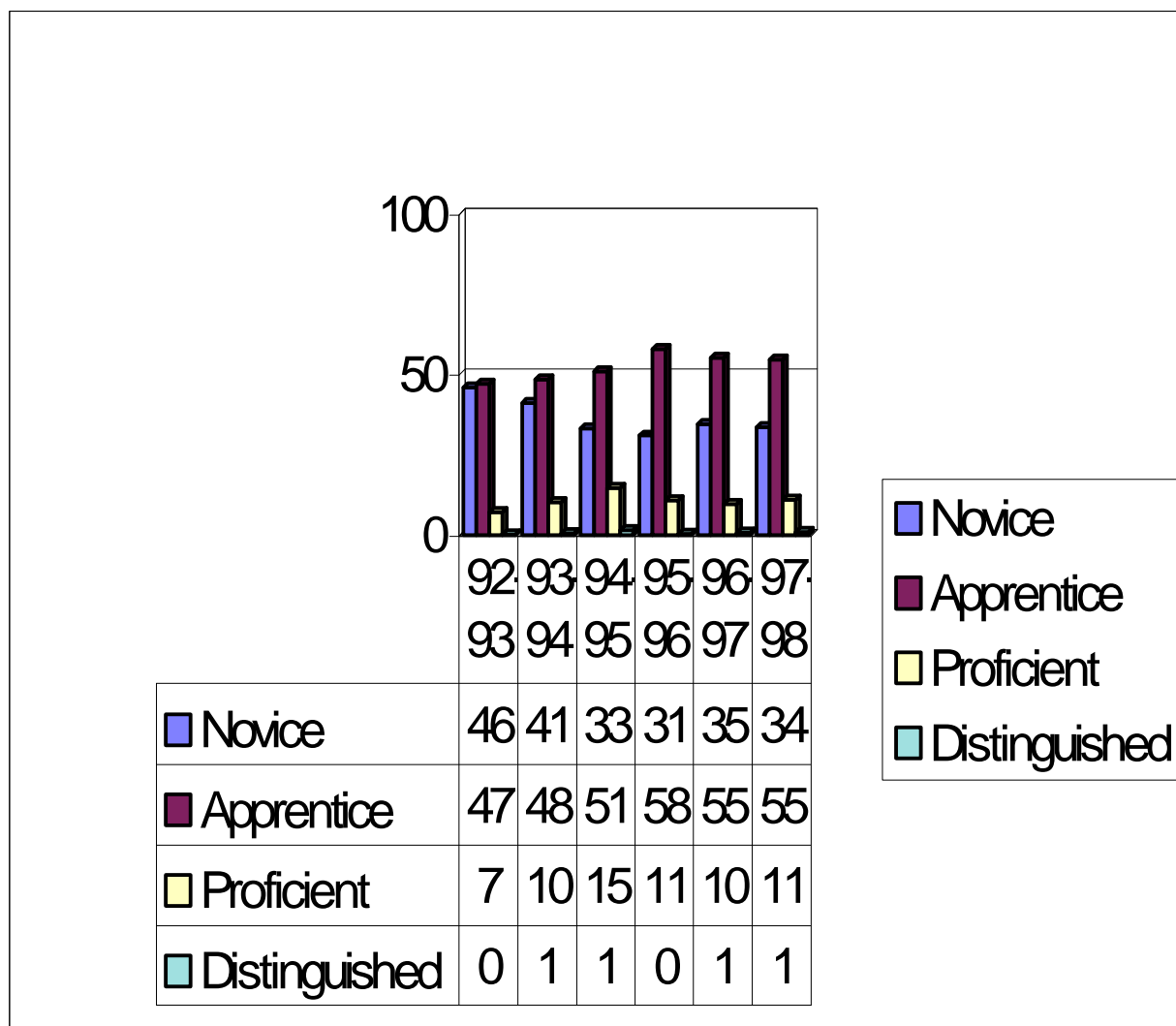


Figure F5. Males' Percentage NAPD in Grade 7/8 Reading, 1993-1998

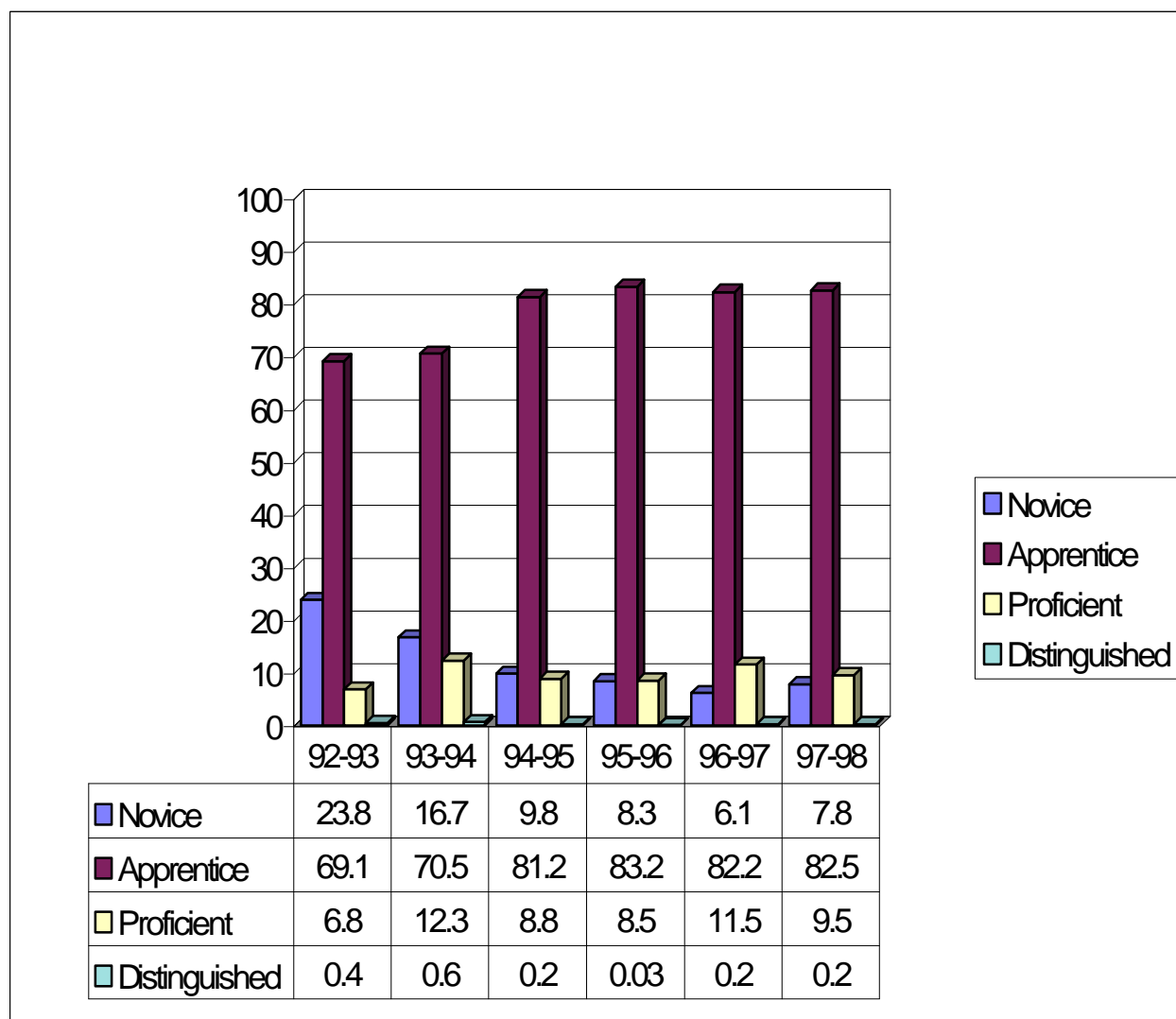


Figure F6. Males' Percentage NAPD in Grades 7 and 8 Math, 1993-1998

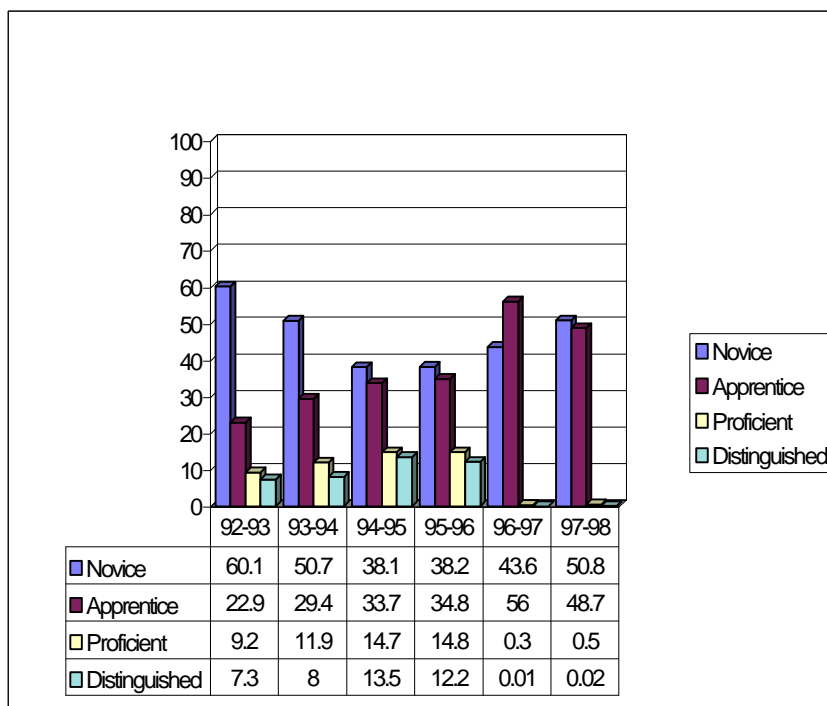


Figure F7. Males' Percentage NAPD in Grade 7/8 Science, 1993-1998

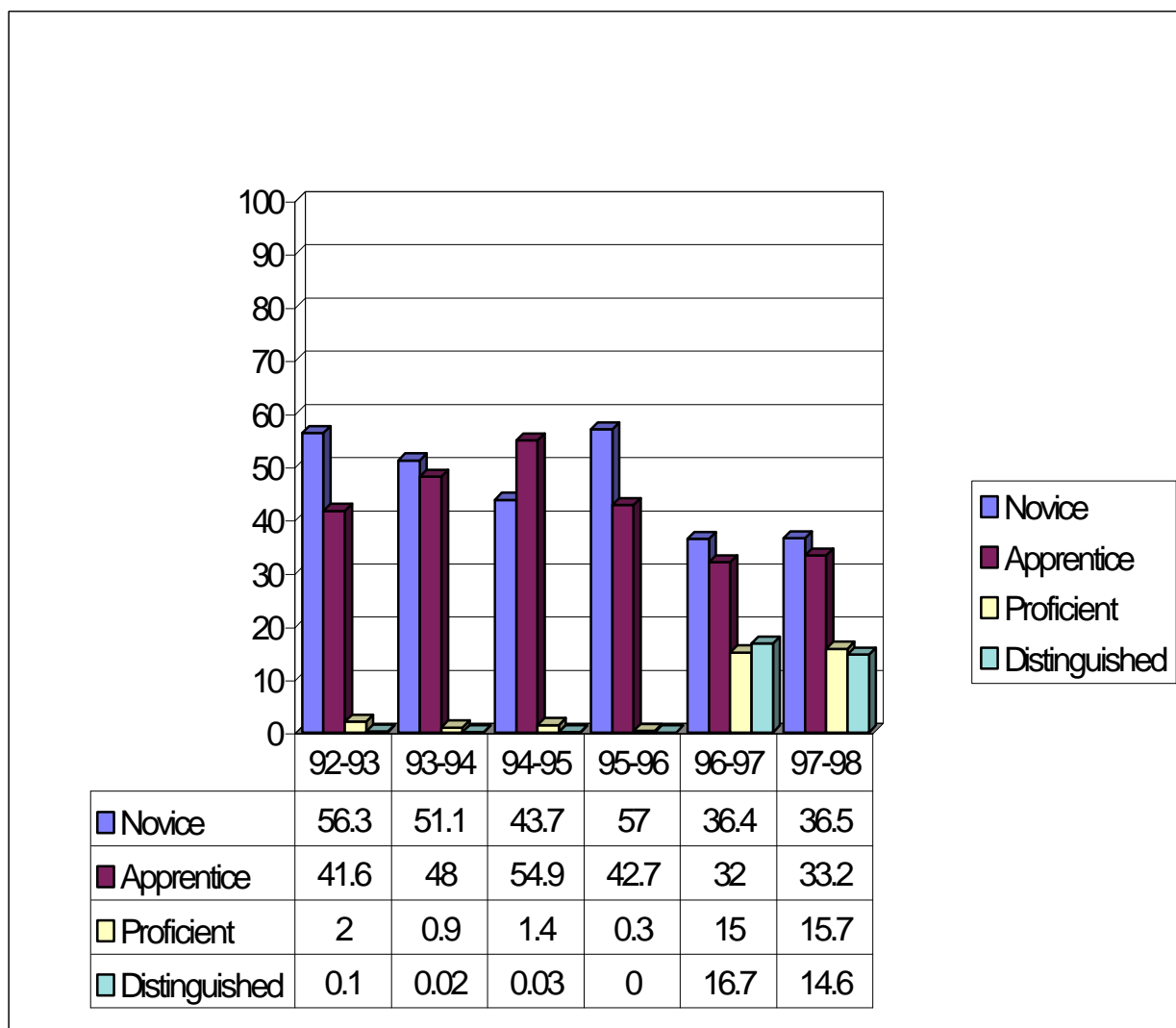


Figure F8. Males' Percentage NAPD in Grade 8 Social Studies, 1993-1998

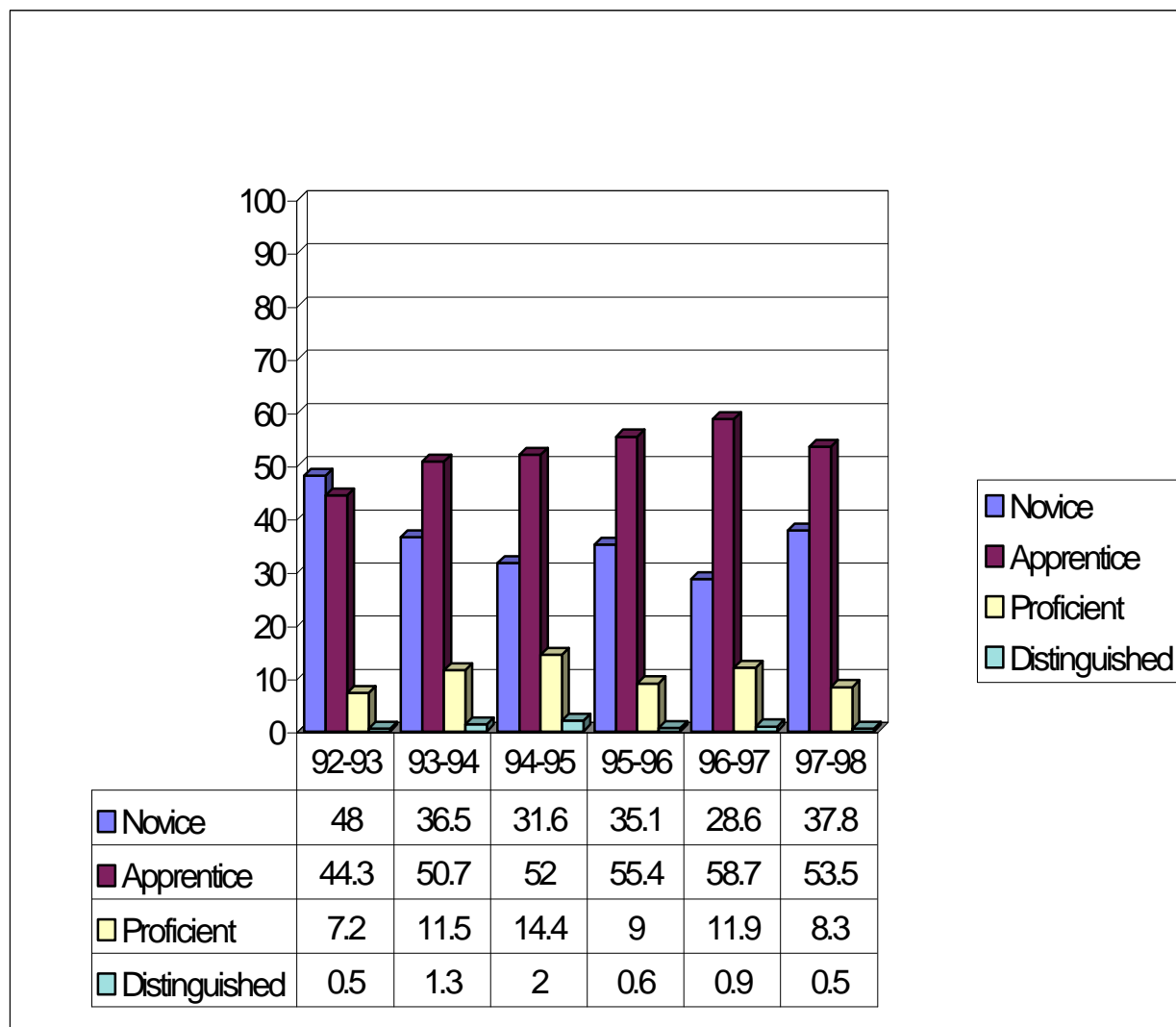


Figure F9. Males' Percentage NAPD in Grade 11/12 Reading, 1993-1998

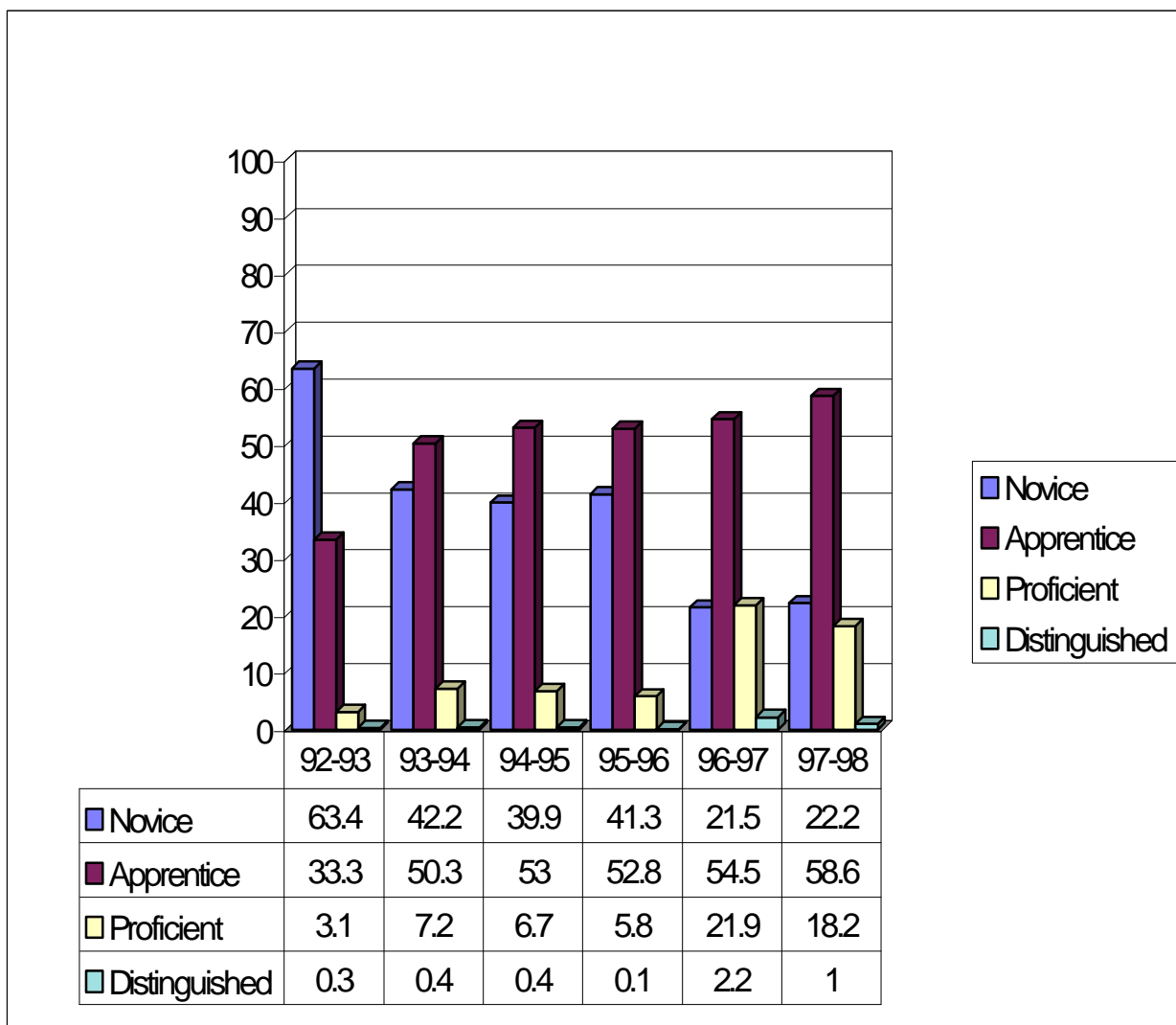


Figure F10. Males' Percentage NAPD in Grade 11/12 Math, 1993-1998

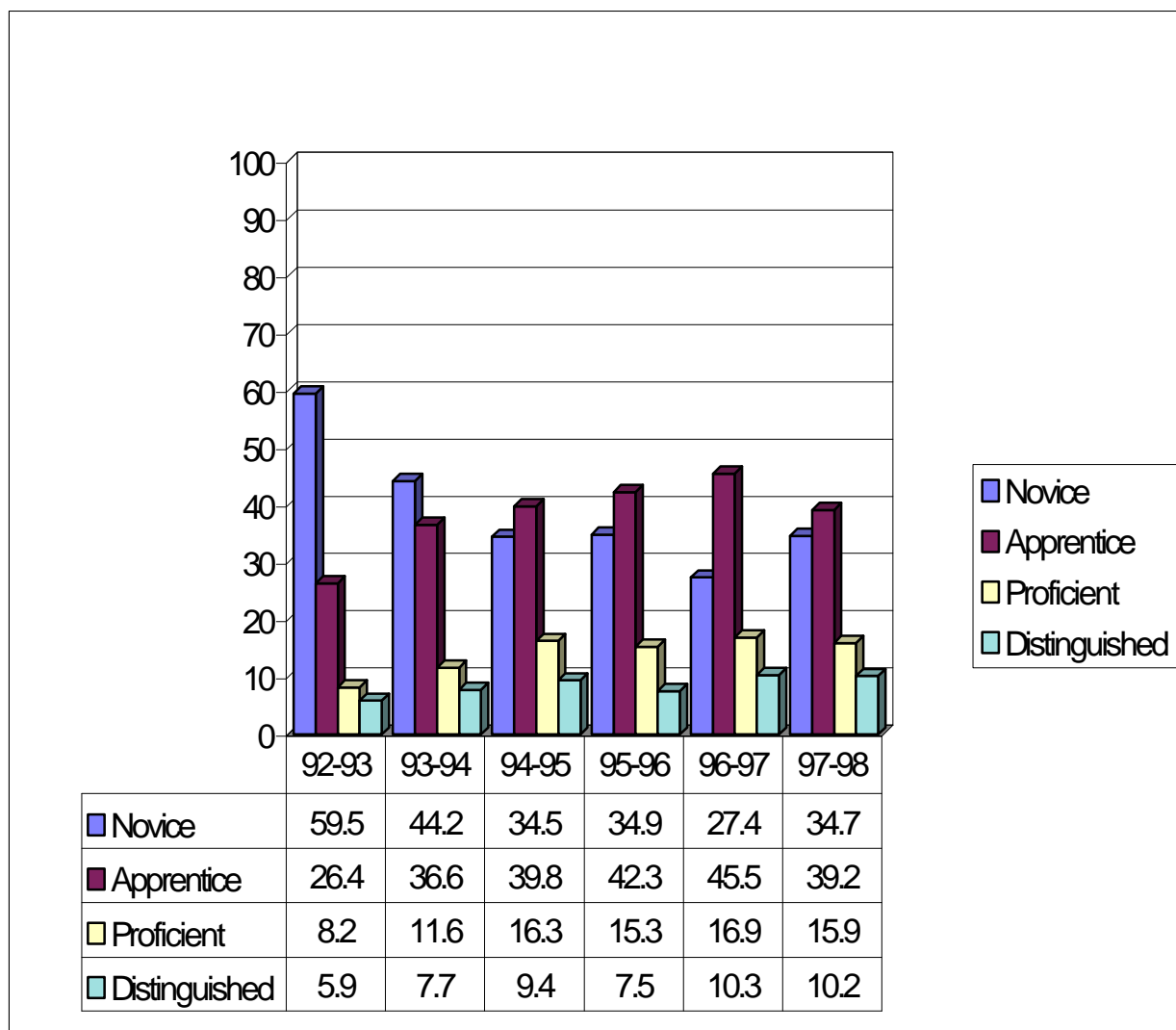


Figure F11. Males' Percentage NAPD in Grade 11/12 Science, 1993-1998

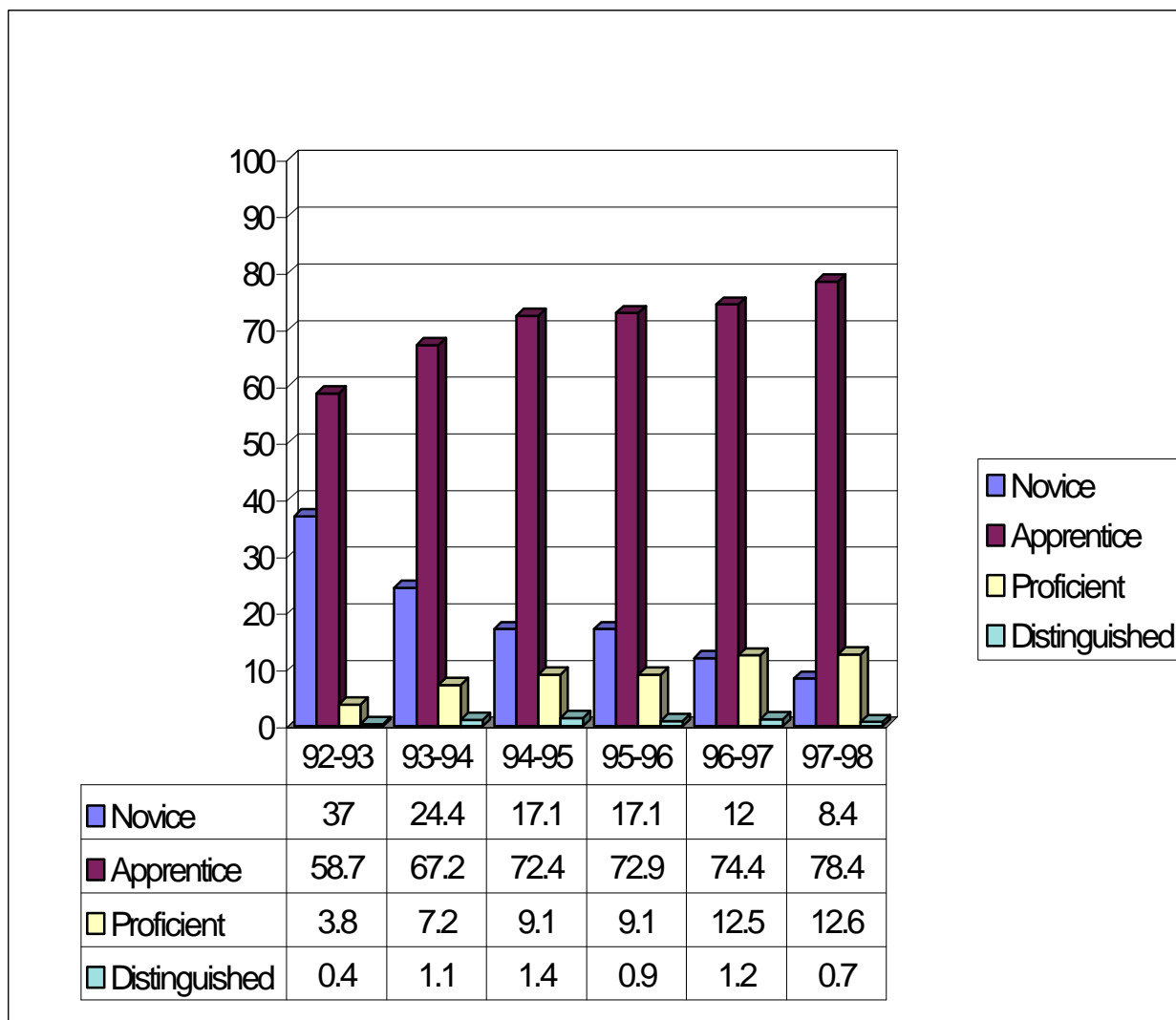
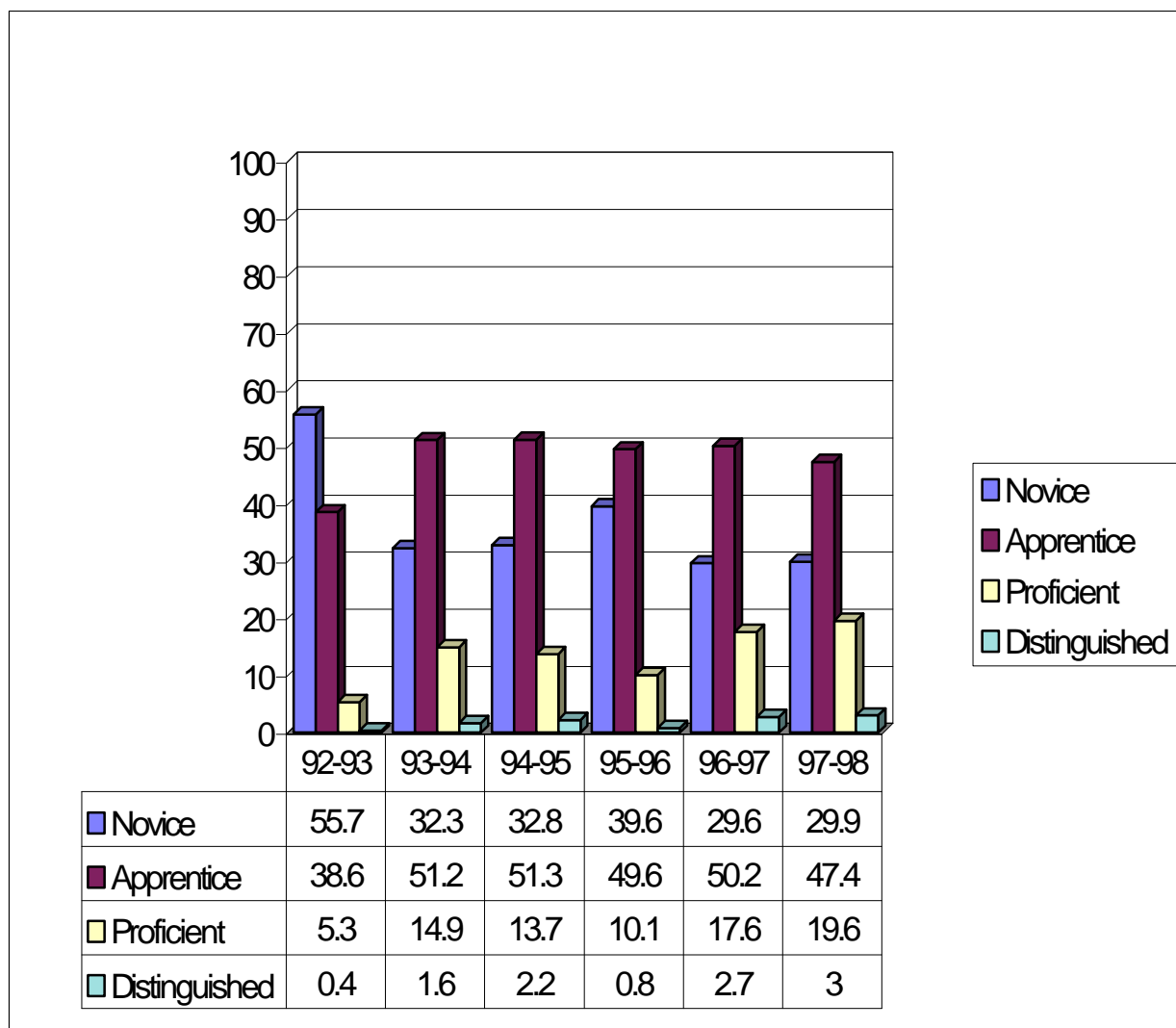


Figure F12. Males' Percentage NAPD in Grade 11/12 Social Studies, 1993-1998



Appendix G⁴⁷

KIRIS Cycle 2 and Cycle 3 Subject Area Performance for
Females by Constructed Response Categories

⁴⁷ Data Tables were added for all Figures in Appendix G by KDE/

Figure G1. Females' Percentage NAPD in Grade 4 Reading, 1993-1998

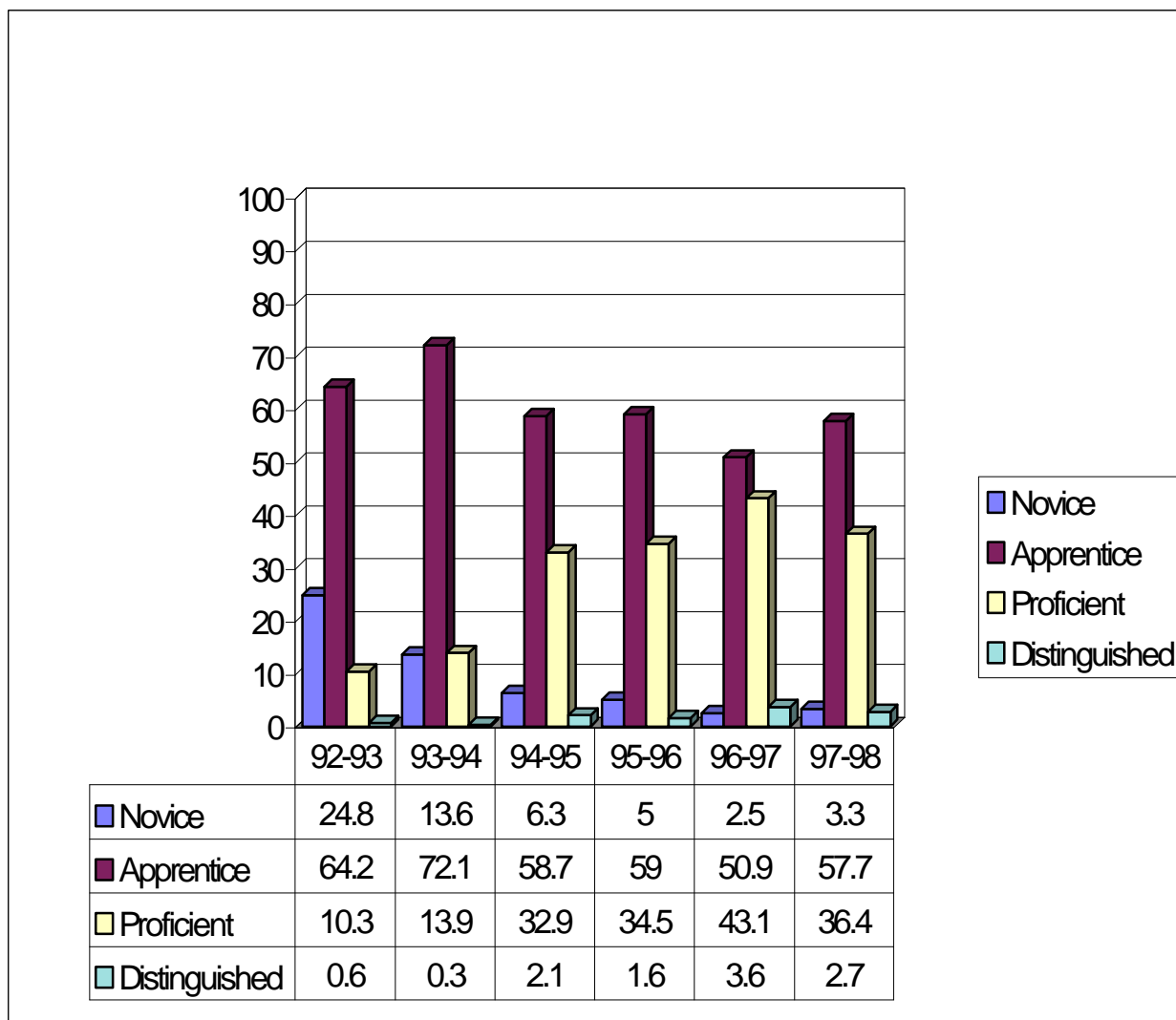


Figure G2. Females' Percentage NAPD in Grade 4/5 Math, 1993-1998

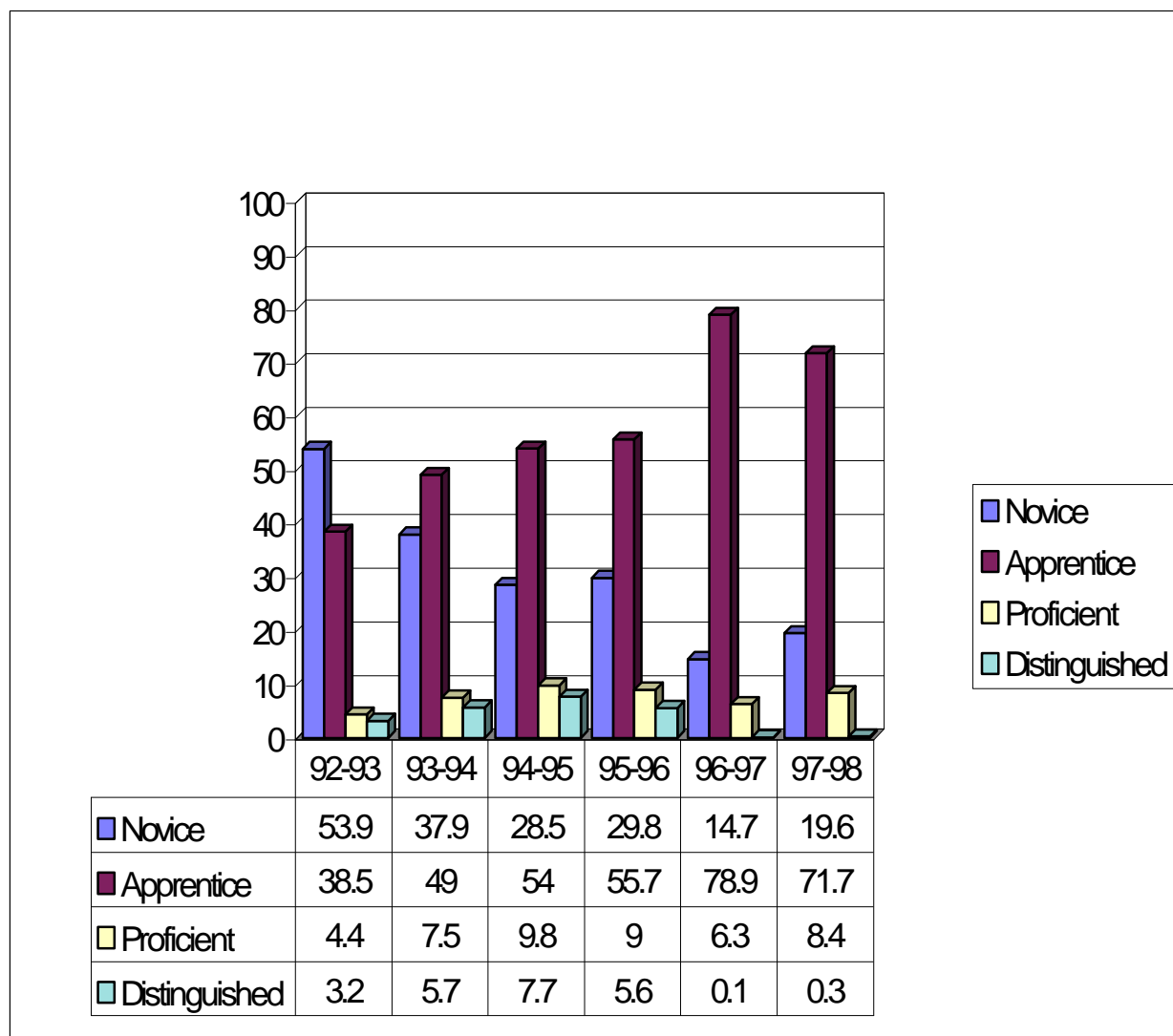


Figure G3 Females' Percentage NAPD in Grade 4 Science, 1993-1998

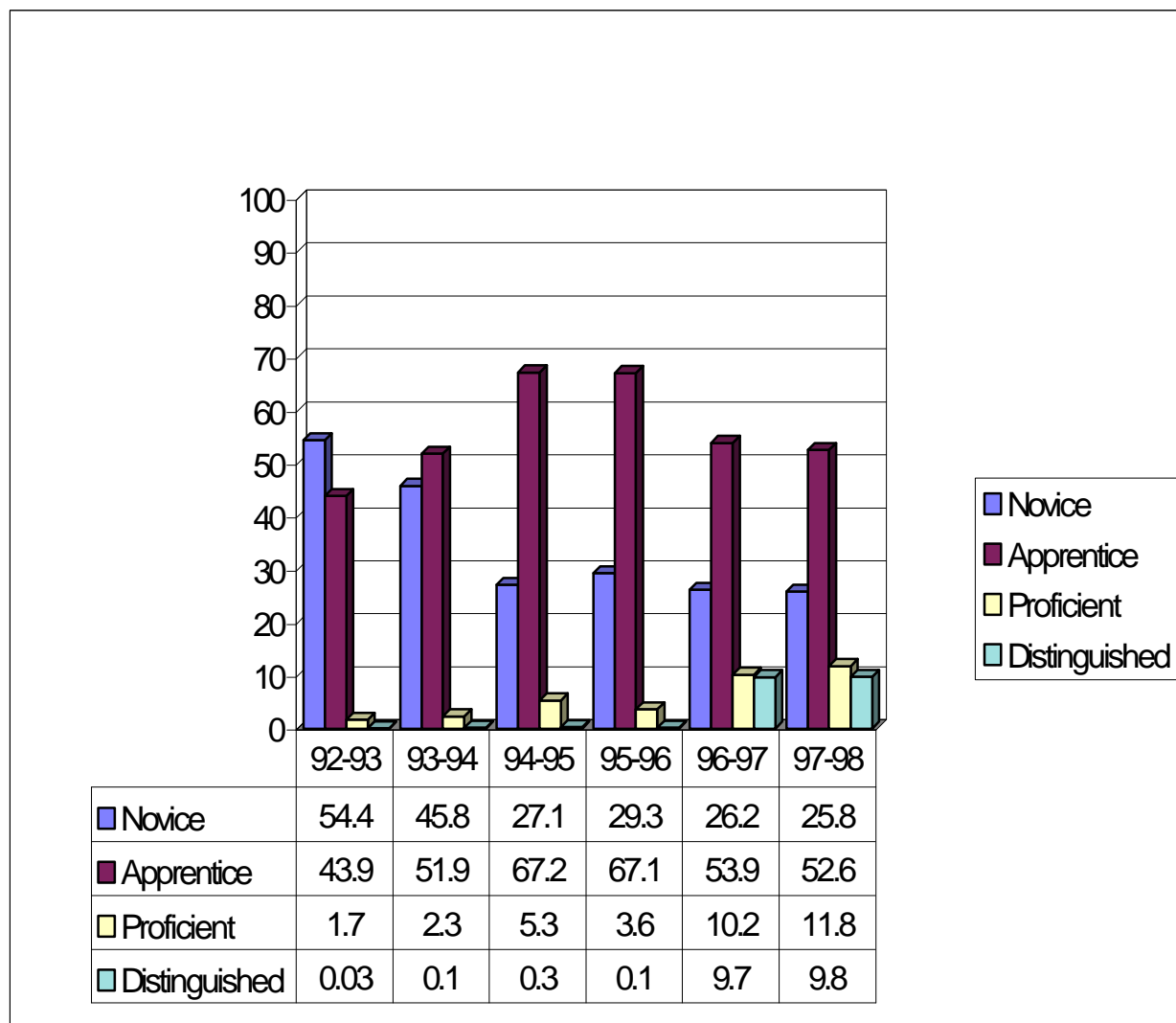


Figure G4. Females' Percentage NAPD in Grade 4/5 Social Studies, 1993-1998.

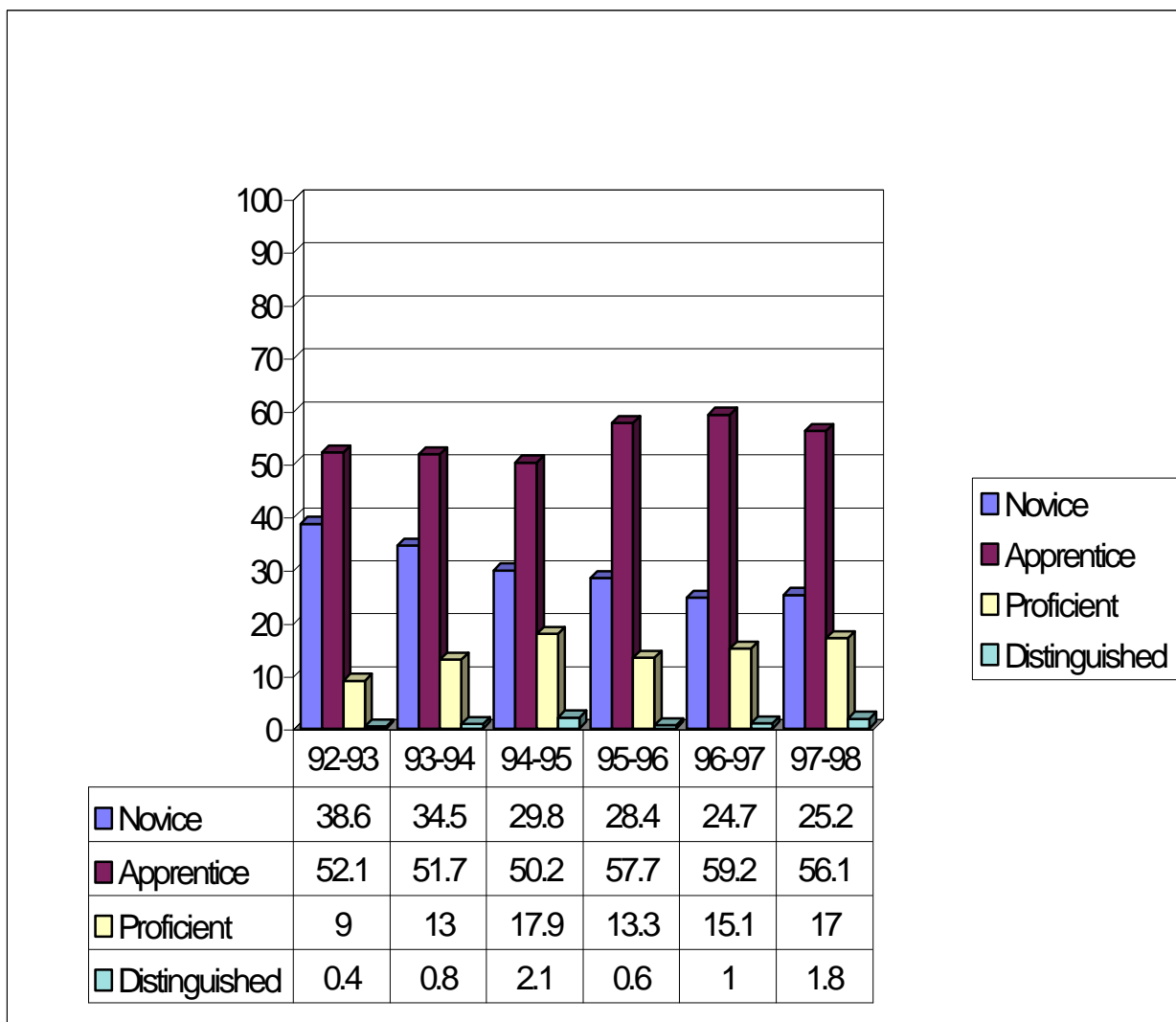


Figure G5. Females' Percentage NAPD in Grade 7/8 Reading, 1993-1998

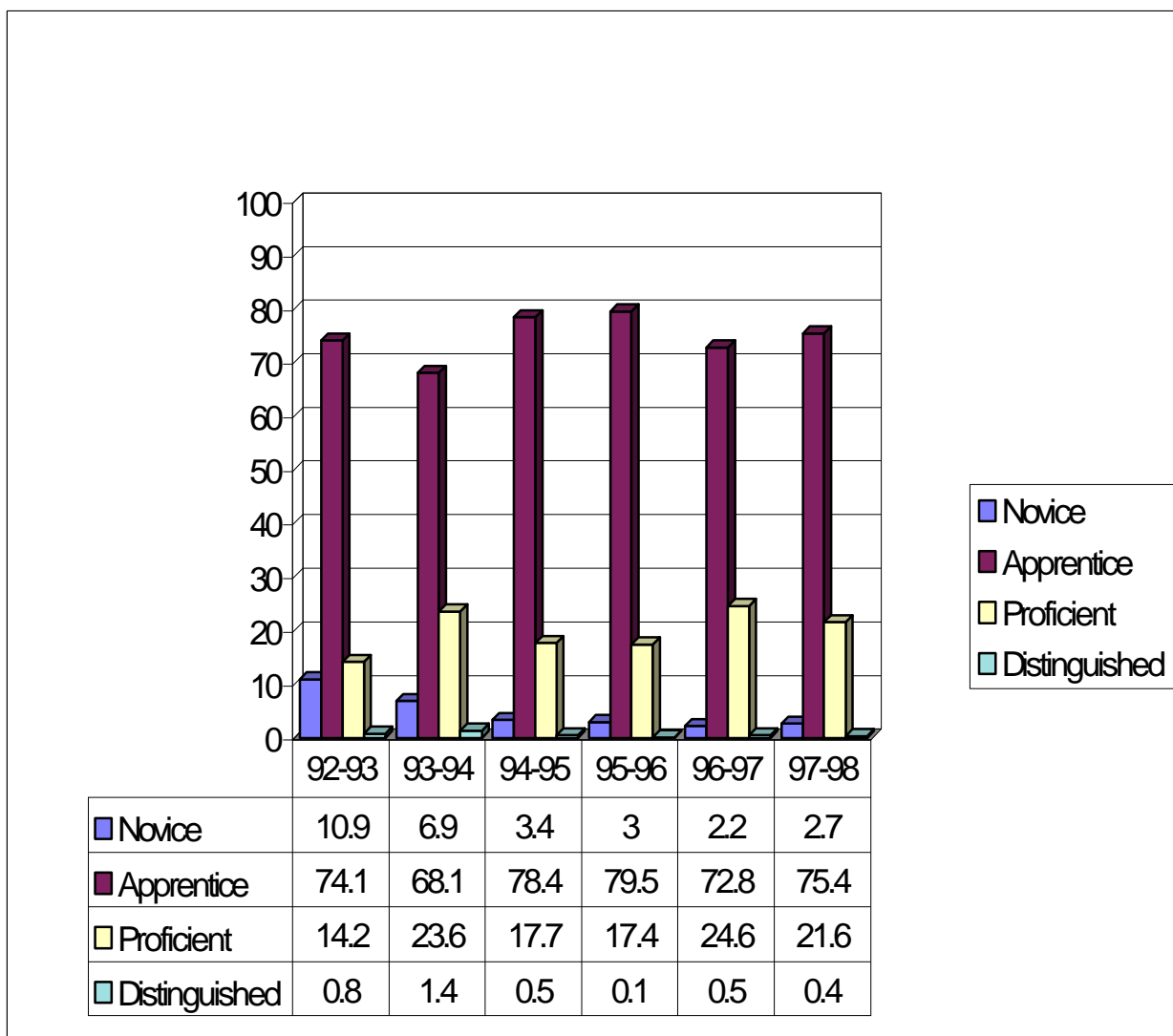


Figure G6. Females' Percentage NAPD in Grade 8 Math, 1993-1998

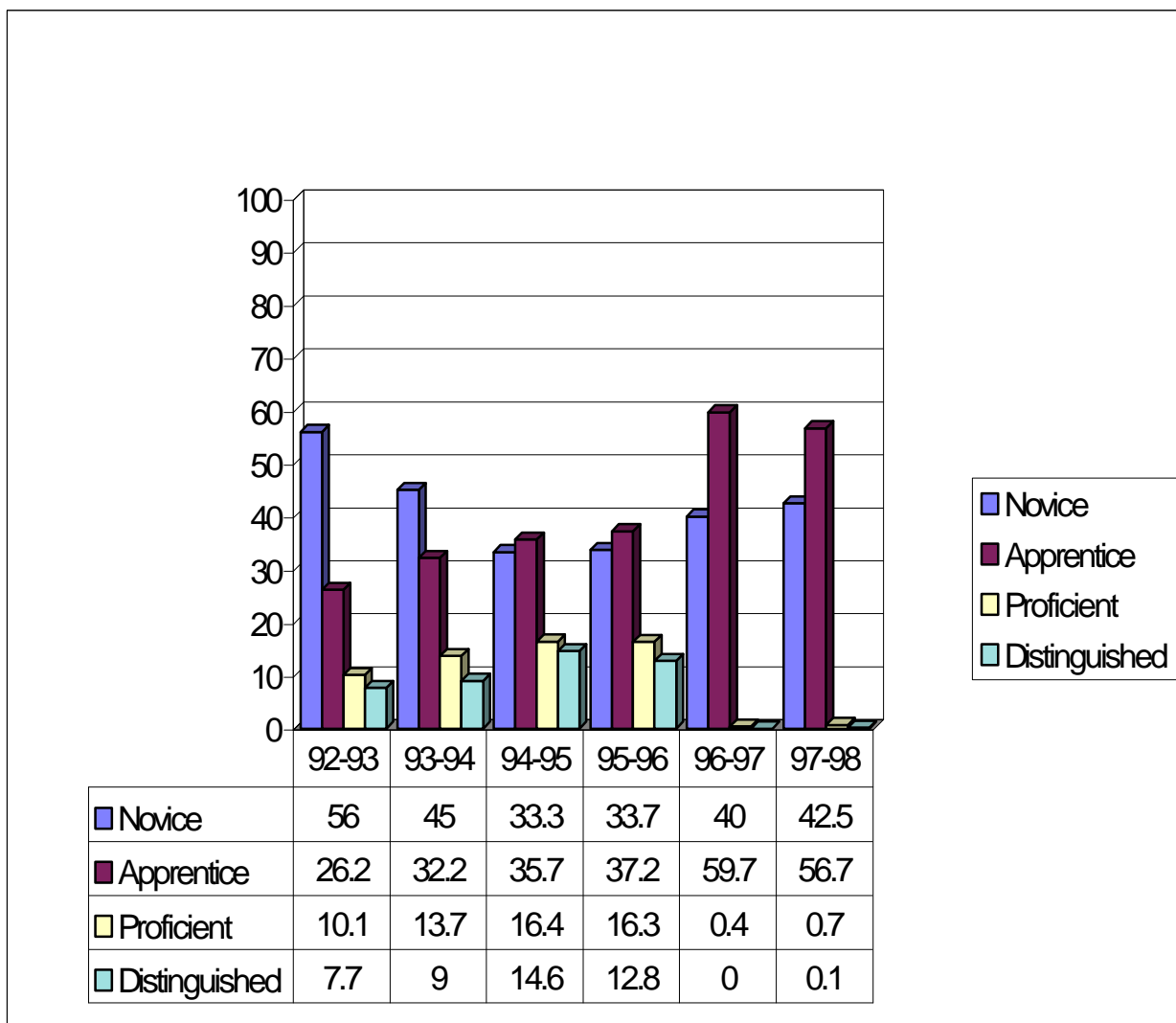


Figure G7. Females' Percentage NAPD in Grade 7/8 Science, 1993-1998

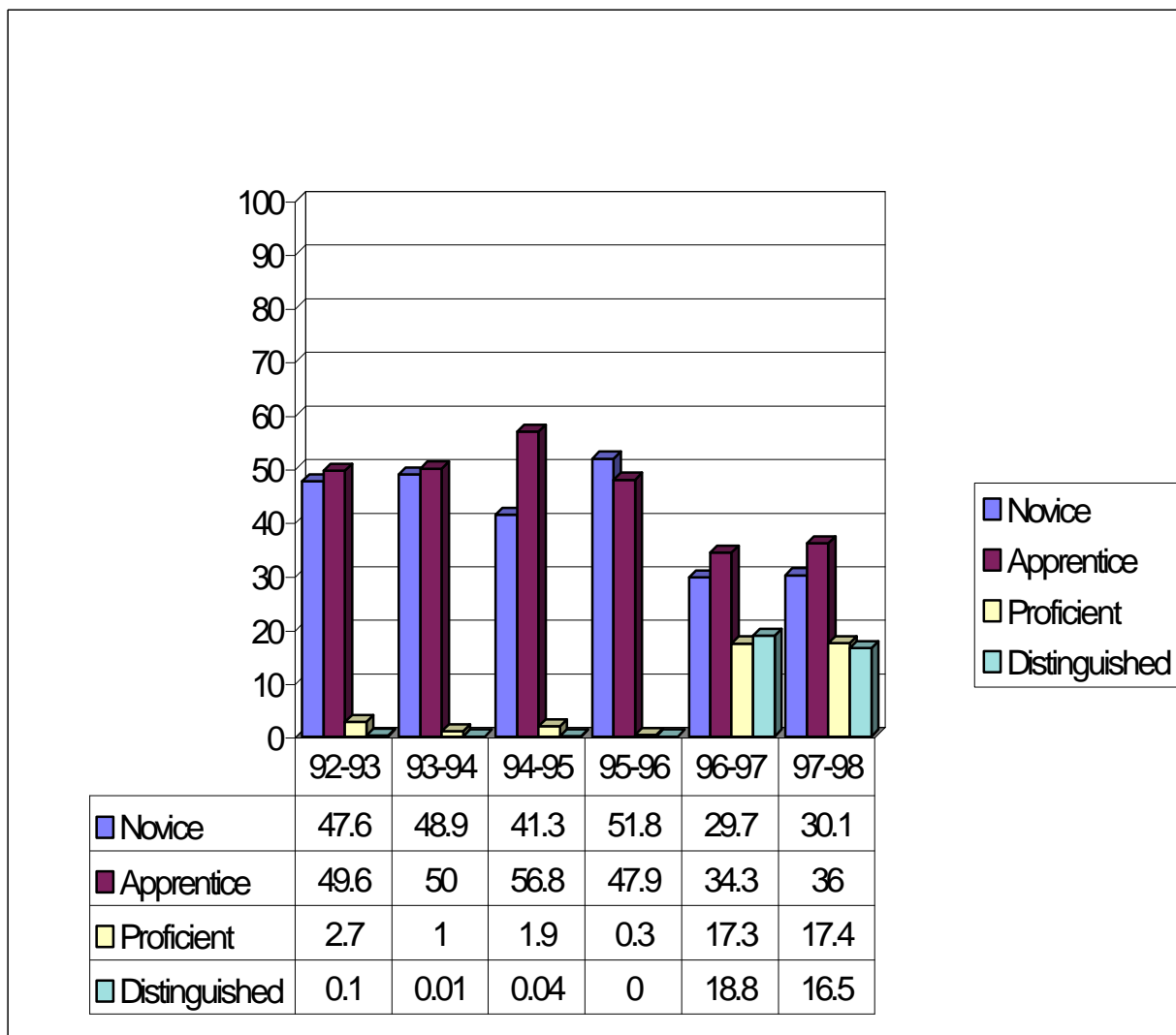


Figure G8. Females' Percentage NAPD in Grade 8 Social Studies, 1993-1998

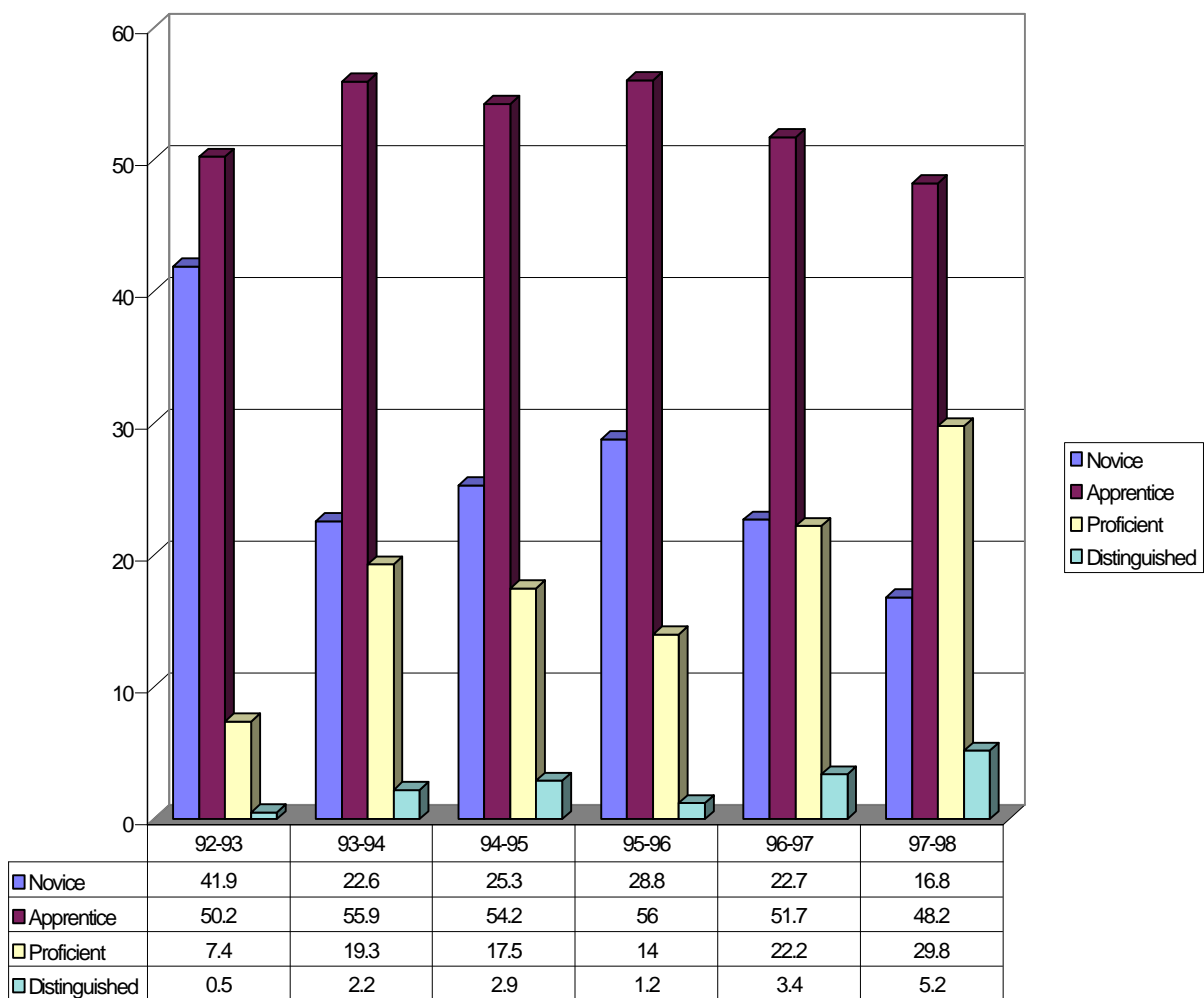


Figure G9. Females' Percentage NAPD in Grade 11/12 Reading, 1993-1998

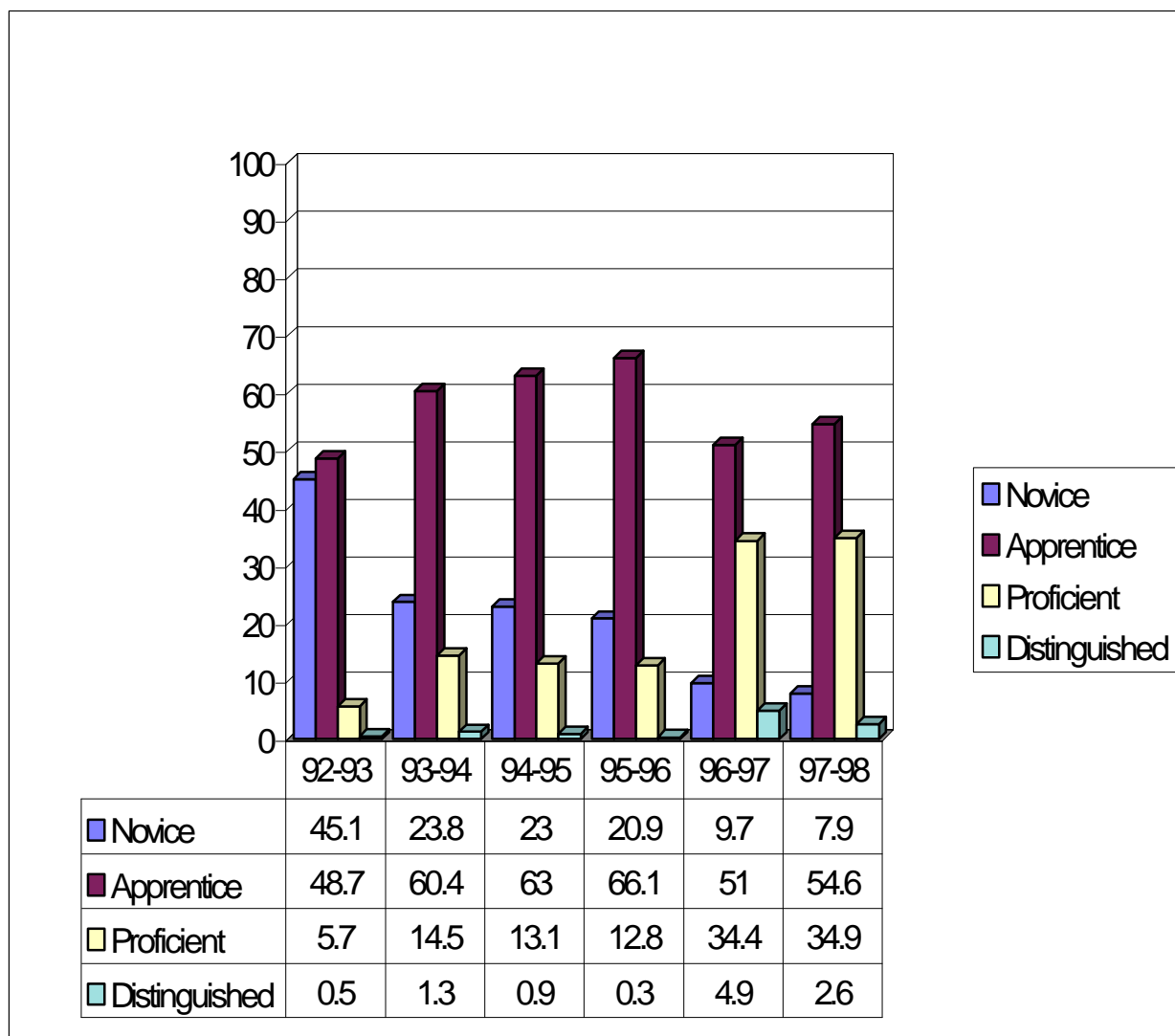


Figure G10. Females' Percentage NAPD in Grade 11/12 Math, 1993-1998

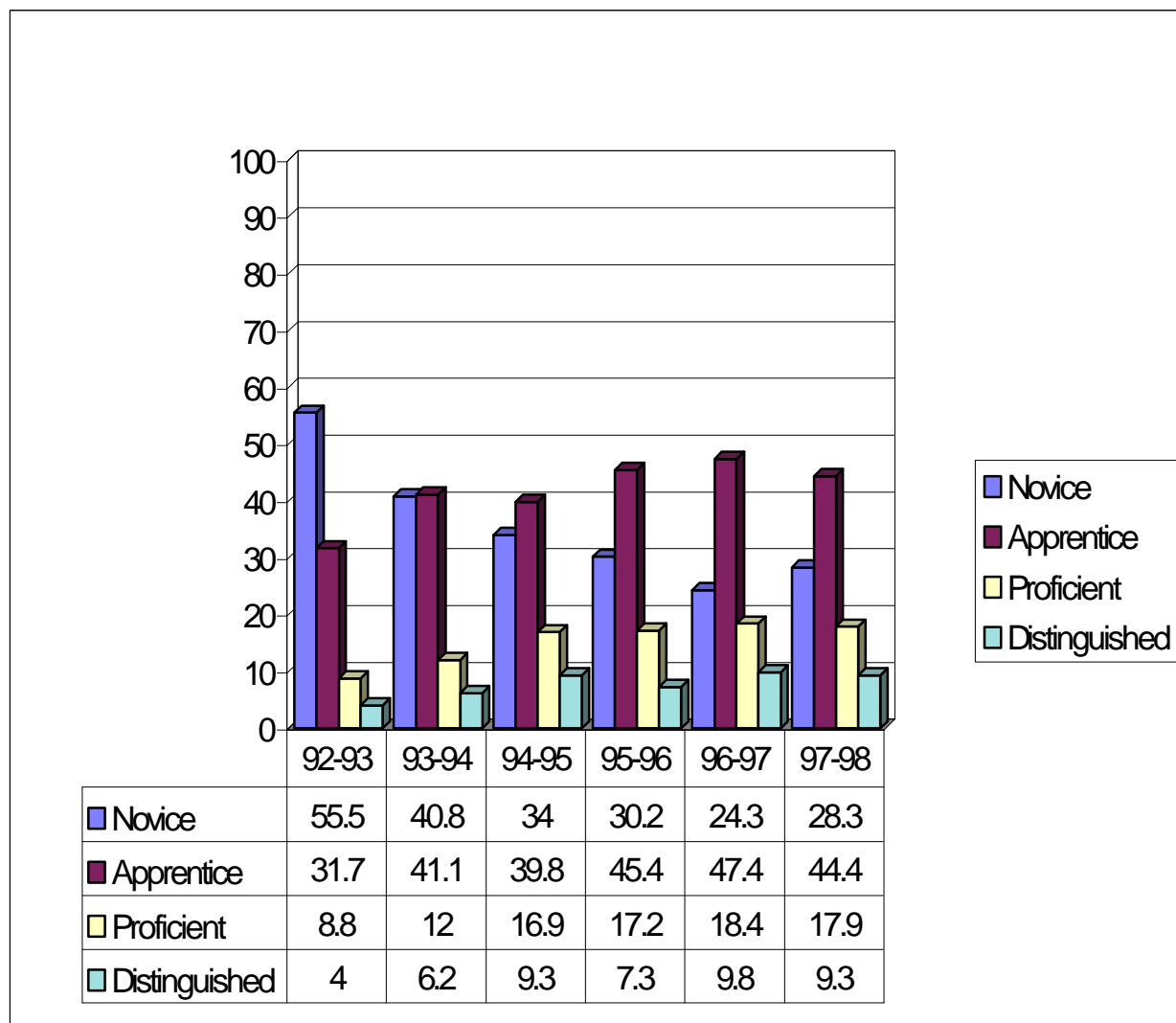


Figure G11. Females' Percentage NAPD in Grade 11/12 Science, 1993-1998

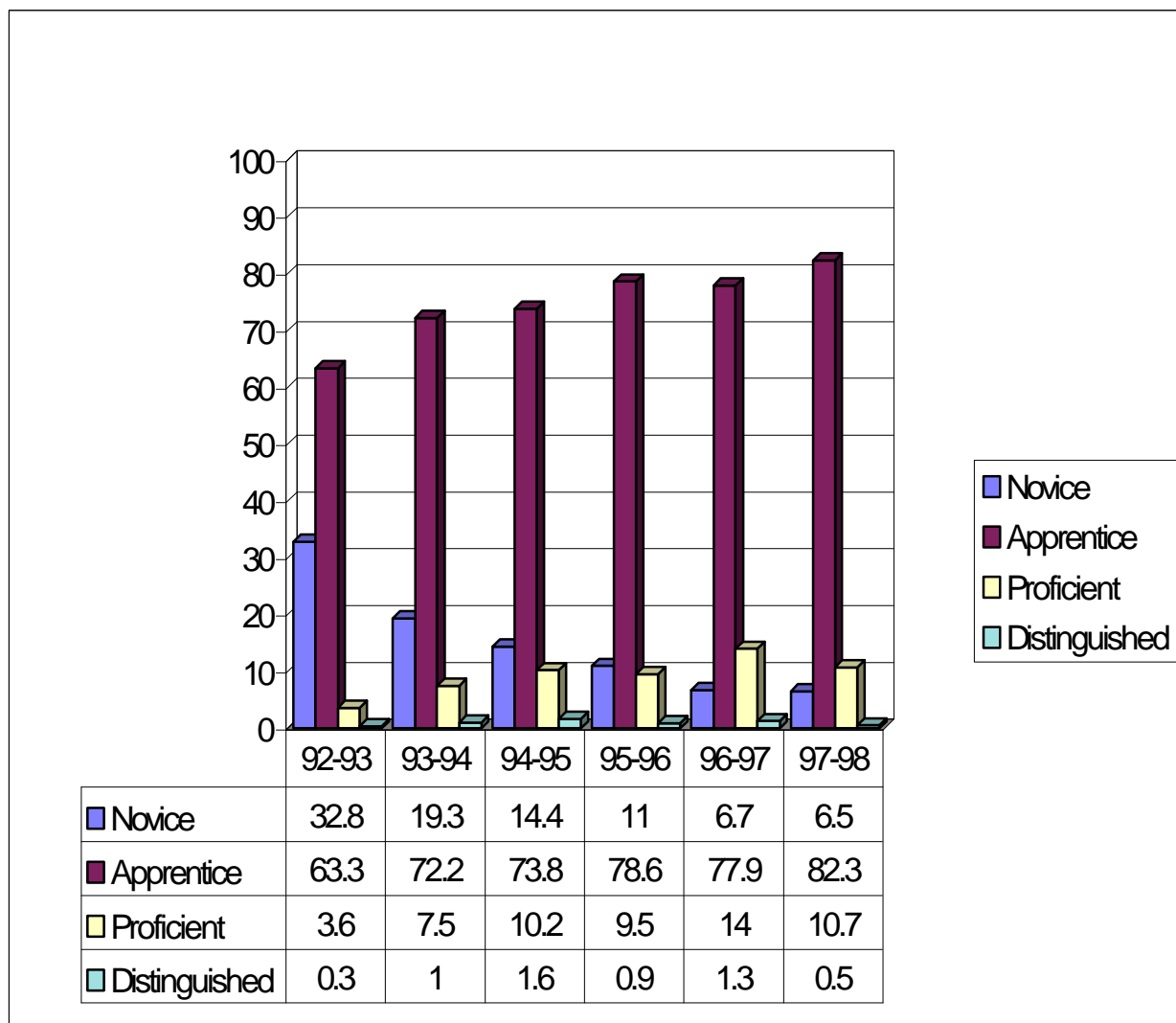
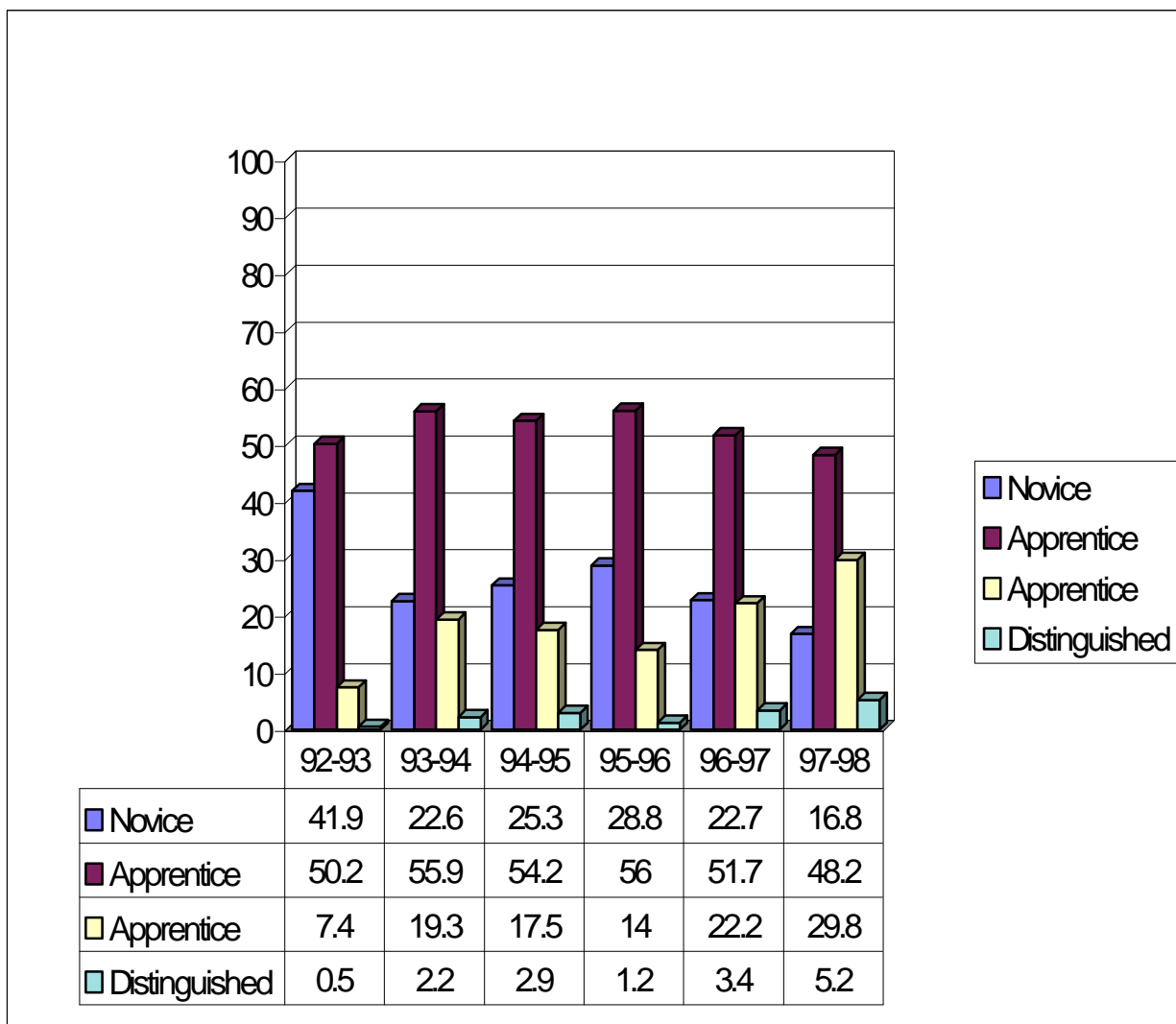


Figure G12. Females' Percentage NAPD in Grade 11/12 Social Studies, 1993-1998



Appendix H⁴⁸

KIRIS Cycle 2 and Cycle 3 Subject Area Performance
for Blacks by Constructed Response Categories

⁴⁸ Data Tables were added for all Figures in Appendix H by KDE.

Figure H1. Blacks' Percentage NAPD in Grade 4 Reading, 1993-1998

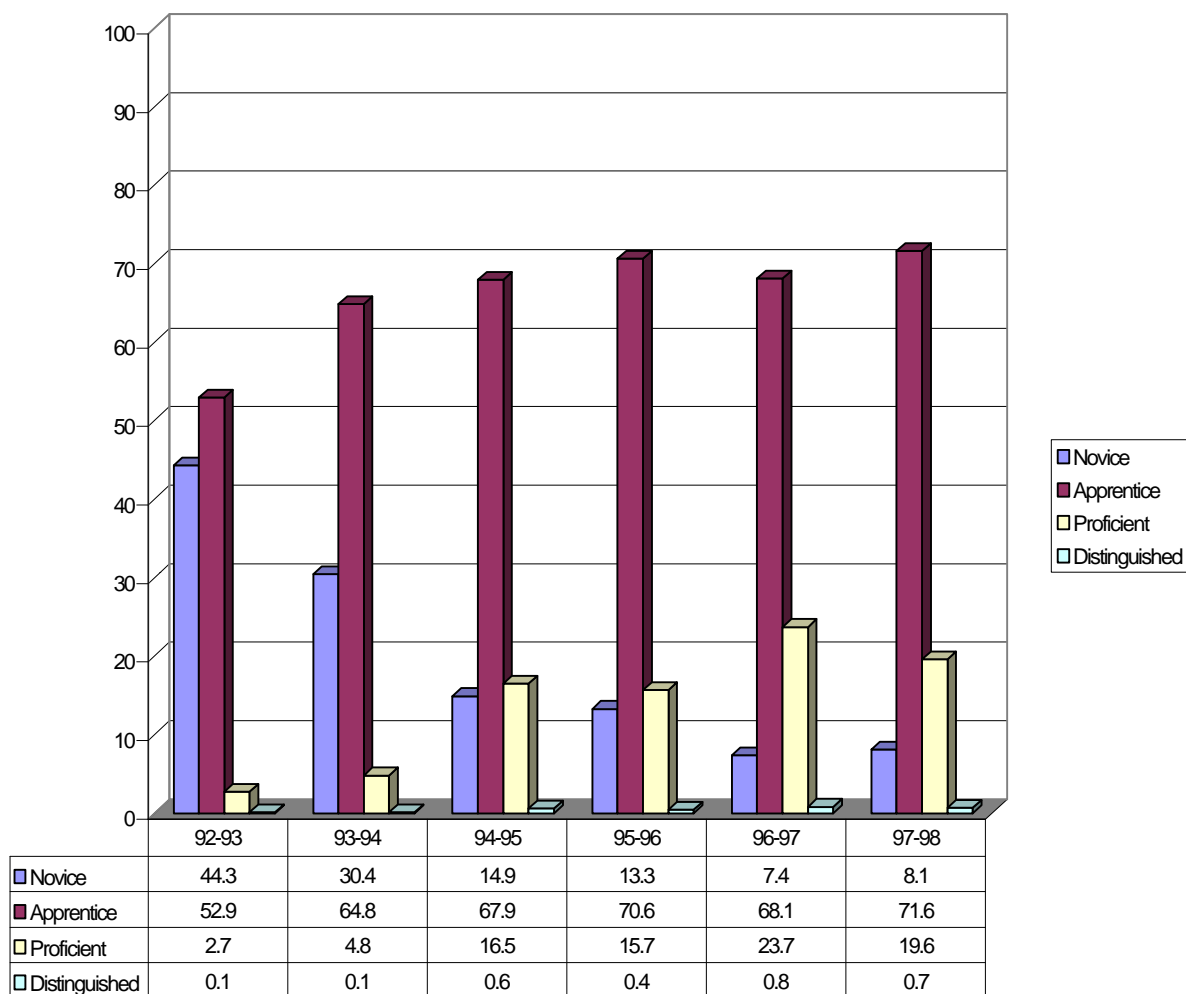


Figure H2. Blacks' Percentage NAPD in Grade 4/5 Math, 1993-1998

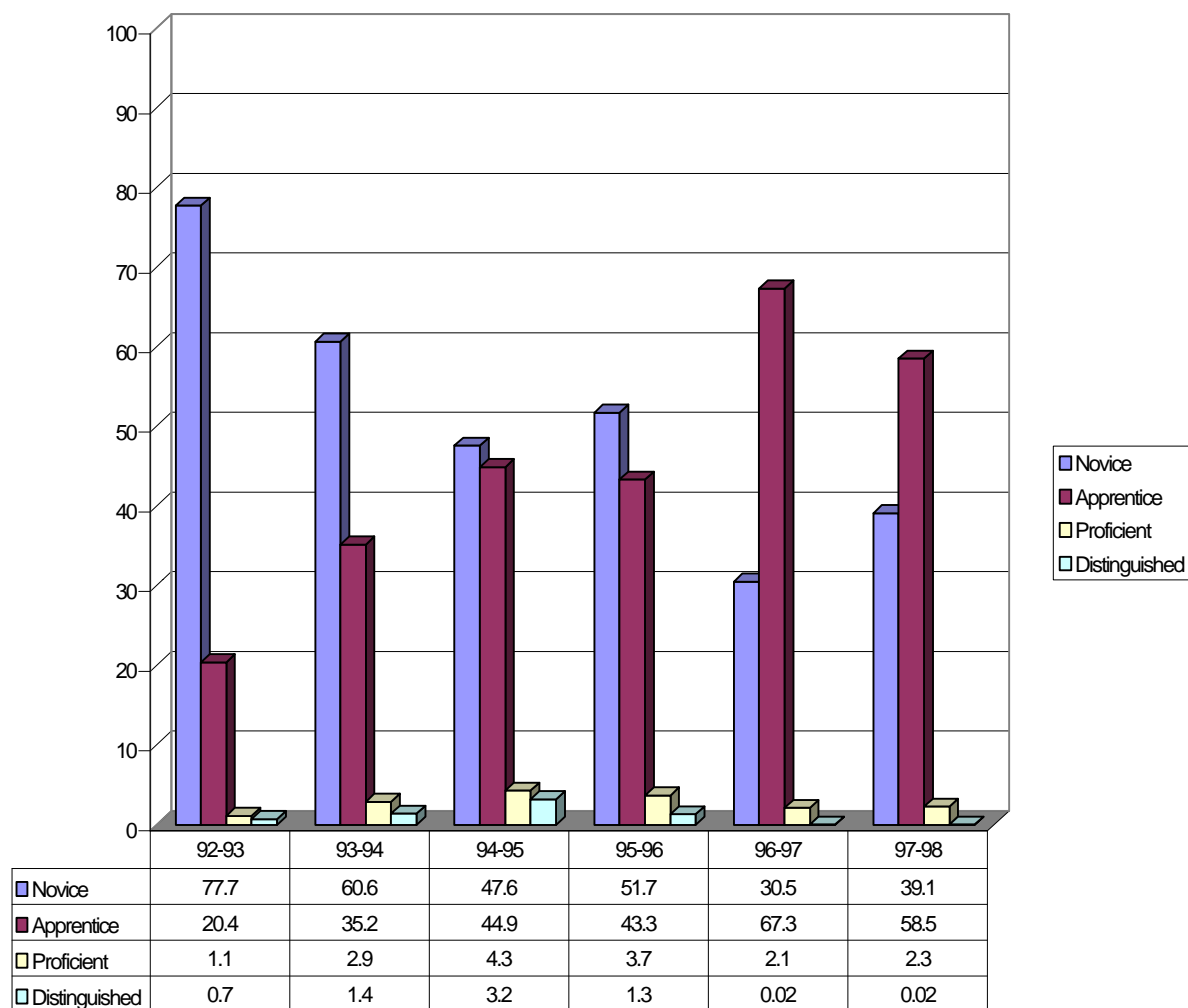


Figure H3. Blacks' Percentage NAPD in Grade 4 Science, 1993-1998

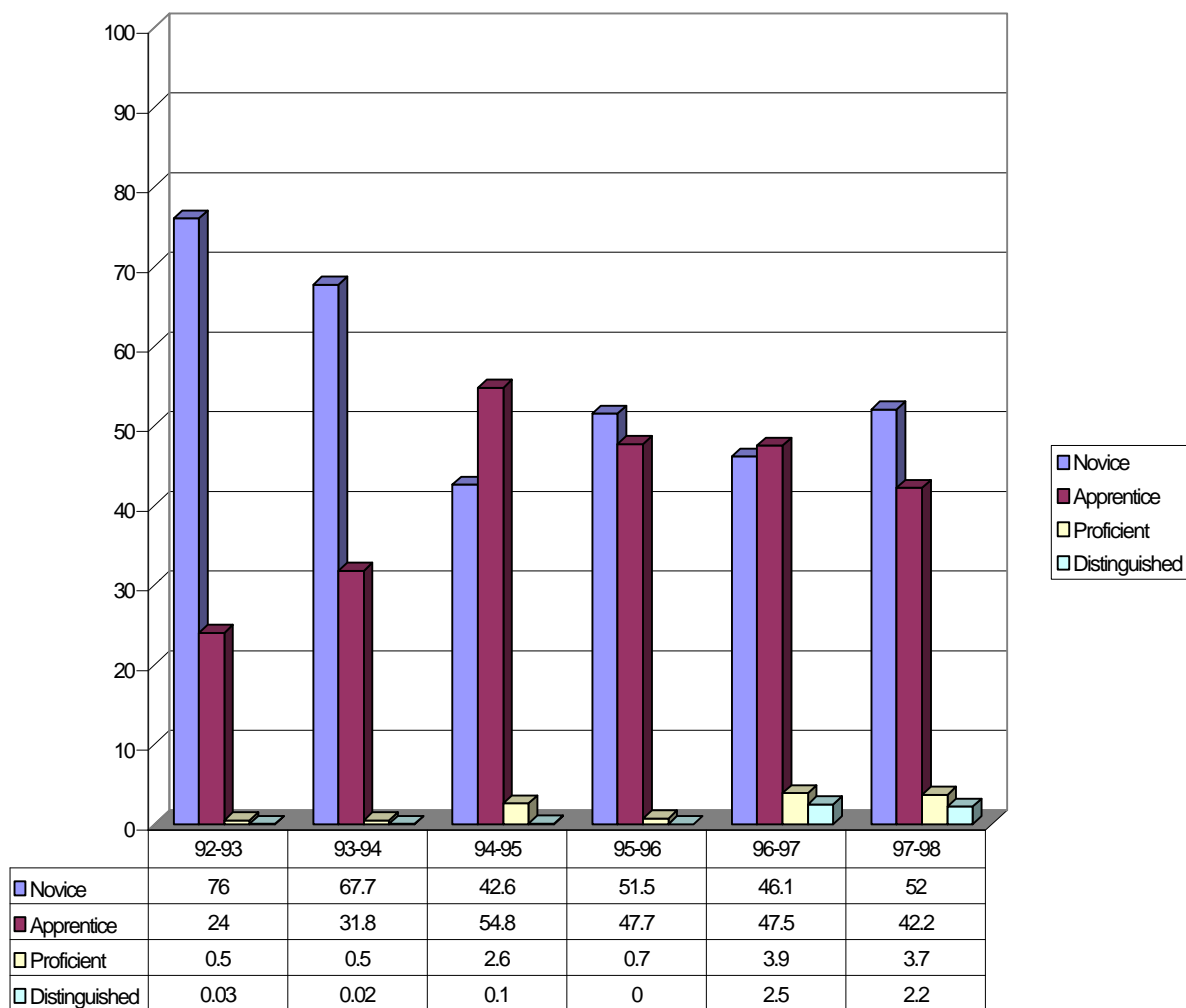


Figure H4. Blacks' Percentage NAPD in Grade 4/5 Social Studies, 1993-1998.

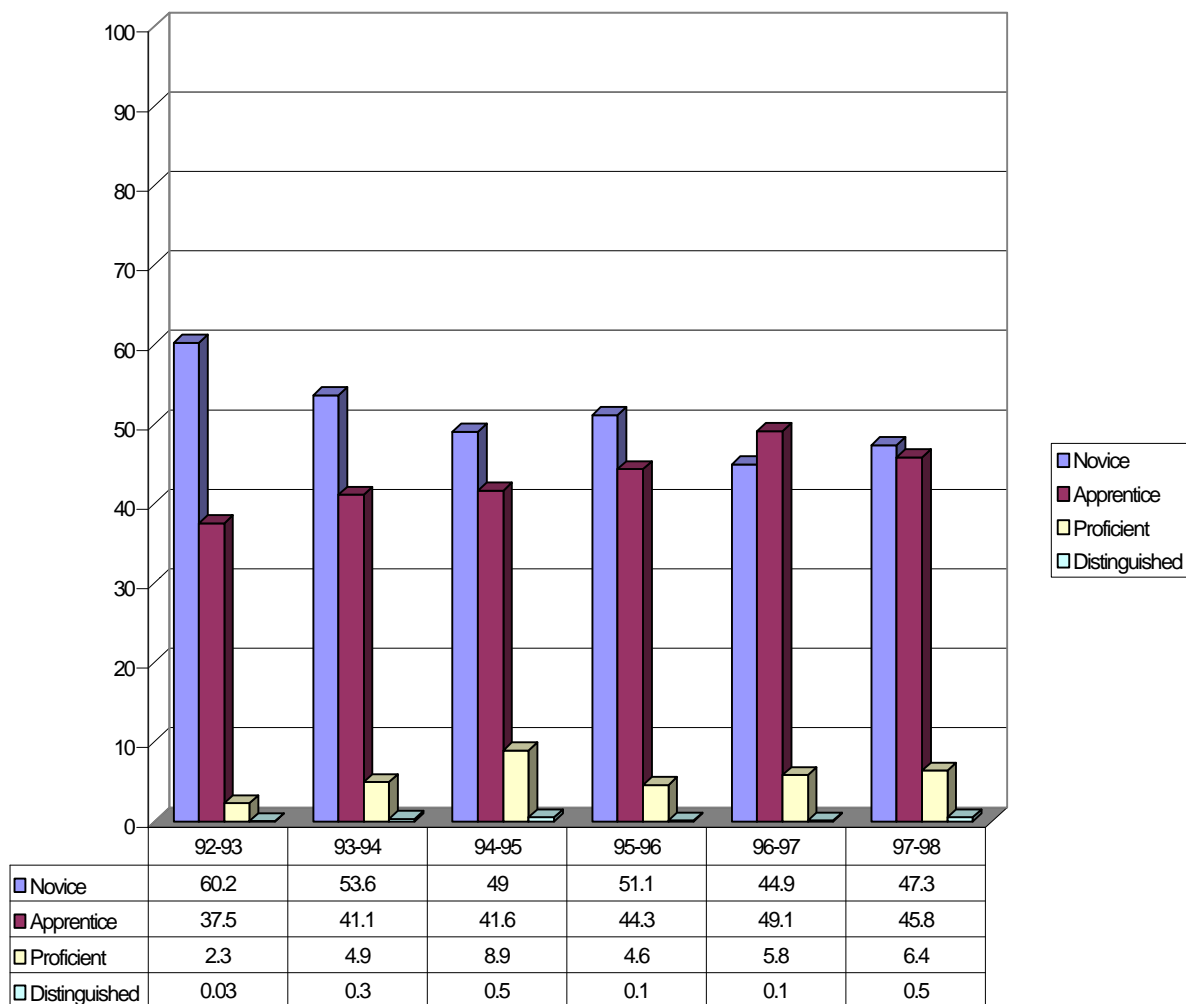


Figure H5. Blacks' Percentage NAPD in Grade 7/8 Reading, 1993-1998

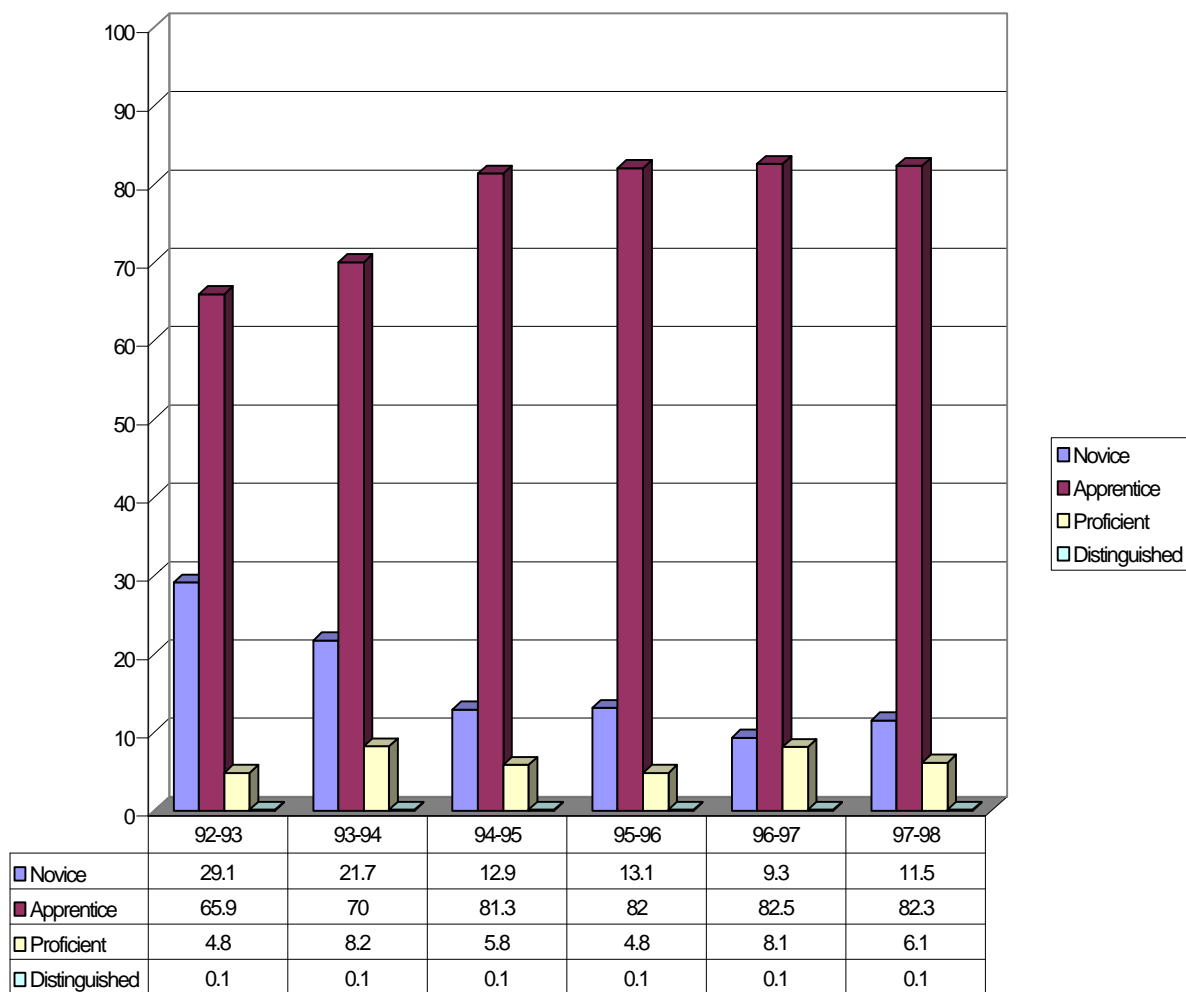


Figure H6. Blacks' Percentage NAPD in Grade 8 Math, 1993-1998

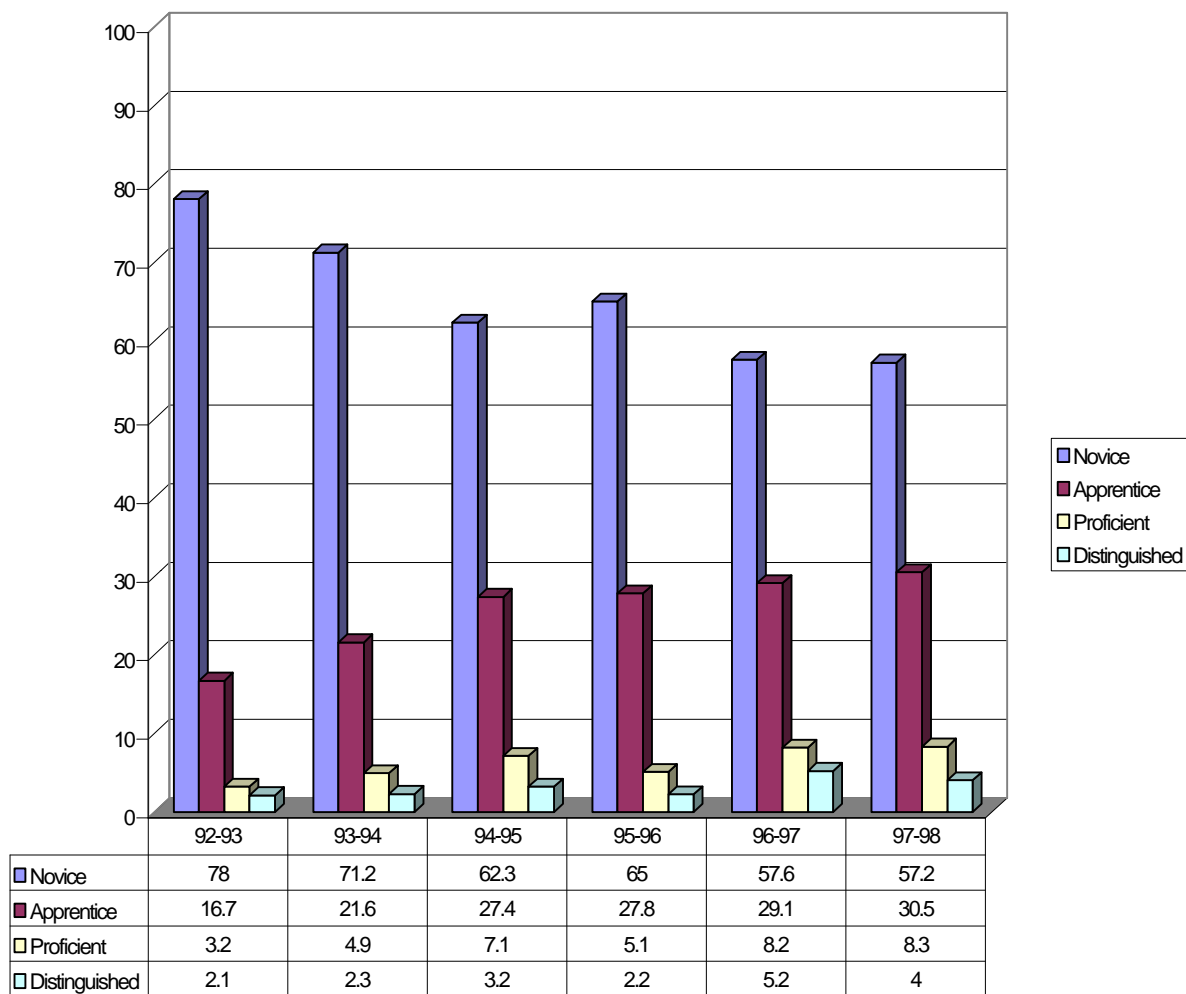


Figure H7. Blacks' Percentage NAPD in Grade 7/8 Science, 1993-1998

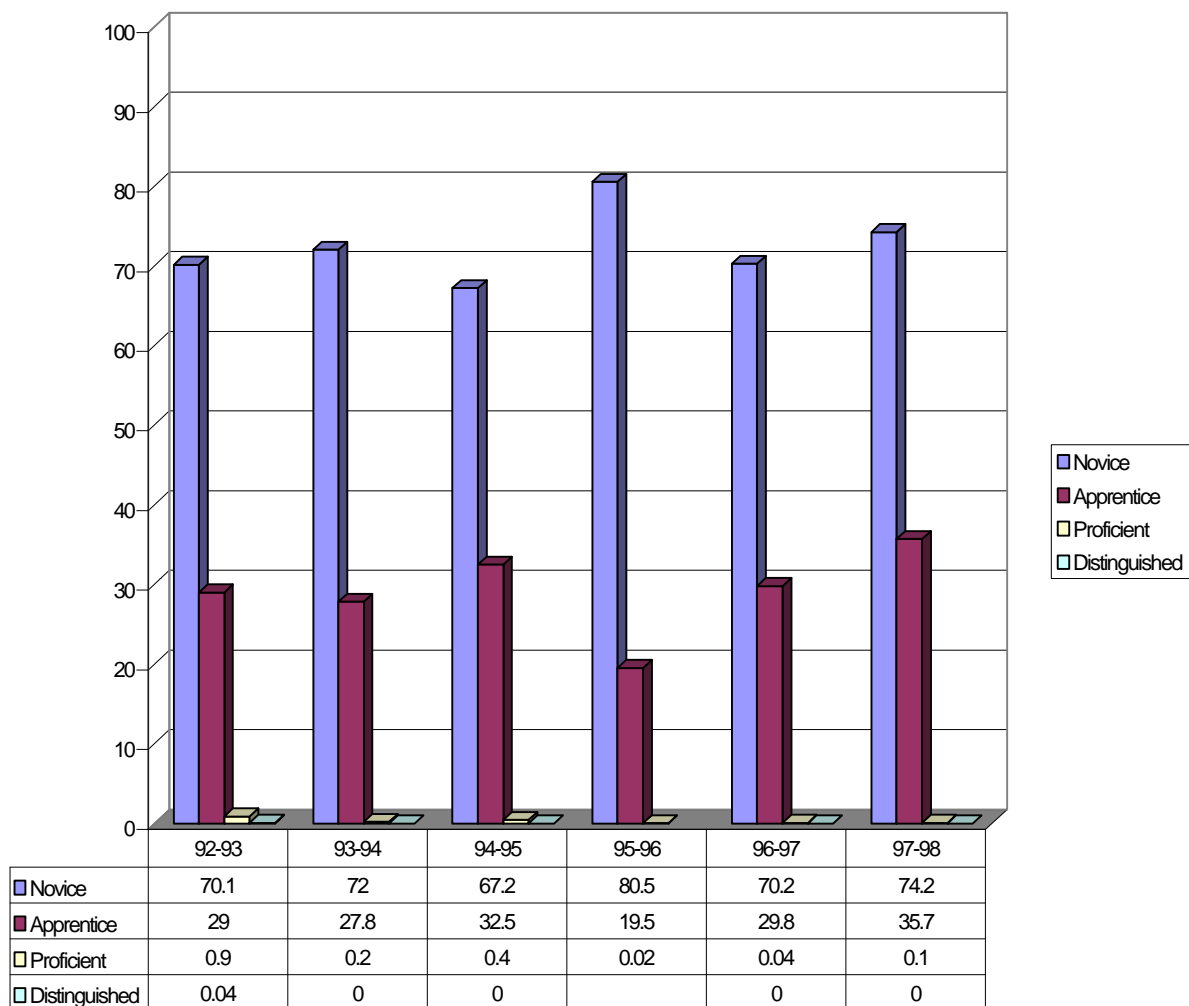


Figure H8. Blacks' Percentage NAPD in Grade 8 Social Studies, 1993-1998

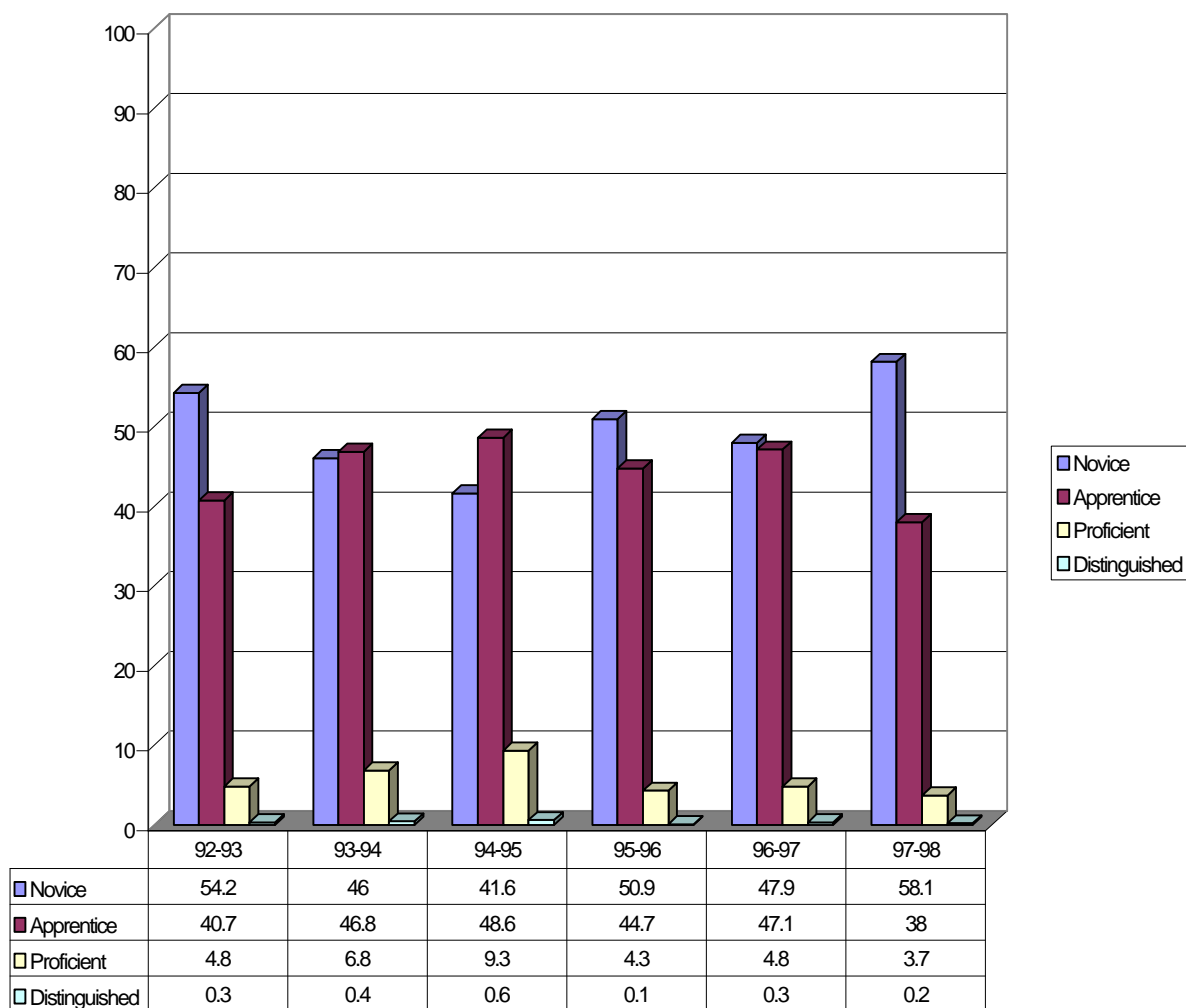


Figure H9. Blacks' Percentage NAPD in Grade 11/12 Reading, 1993-1998

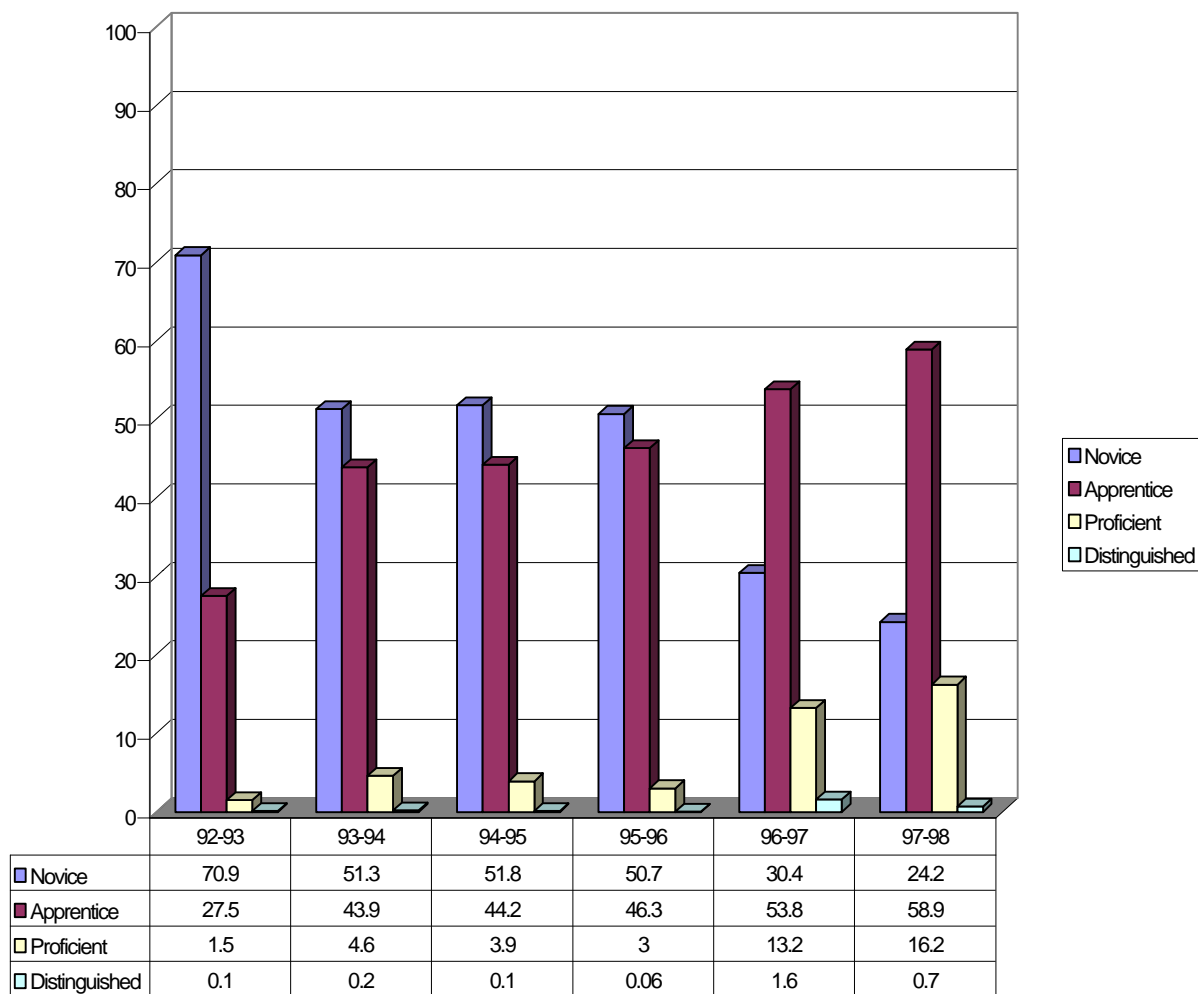


Figure H10. Blacks' Percentage NAPD in Grade 11/12 Math, 1993-1998

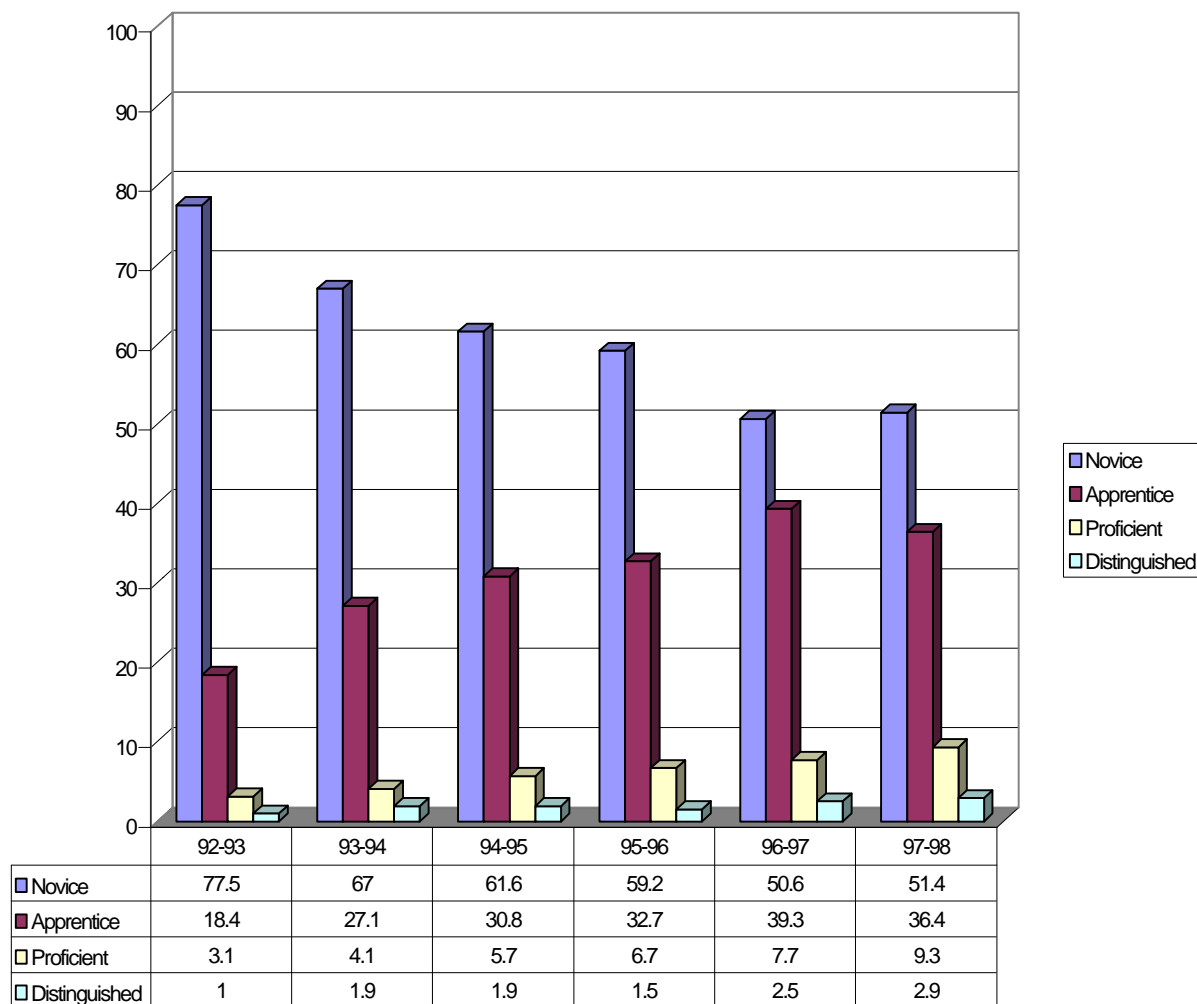


Figure H11. Blacks' Percentage NAPD in Grade 11/12 Science, 1992-1998

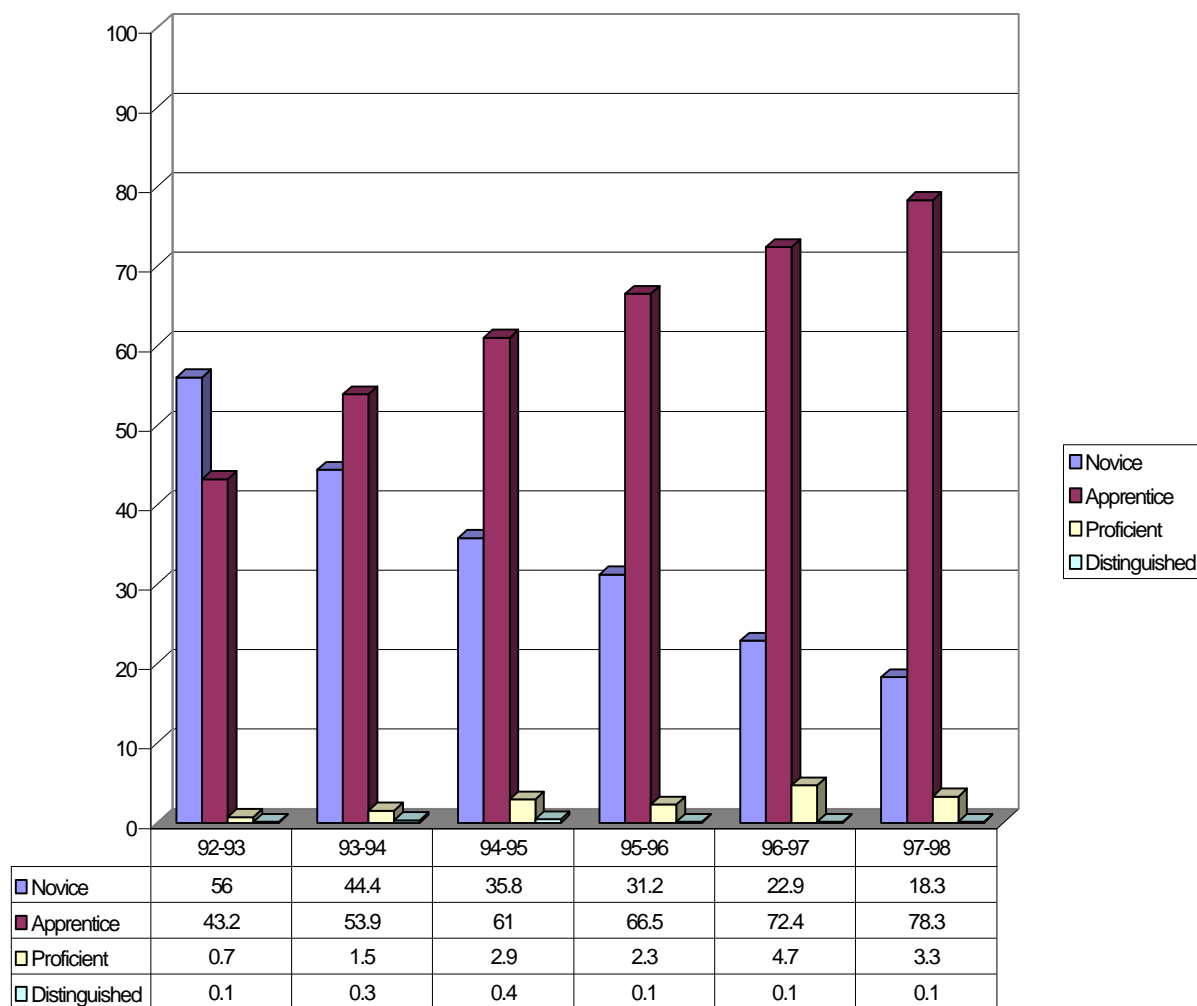
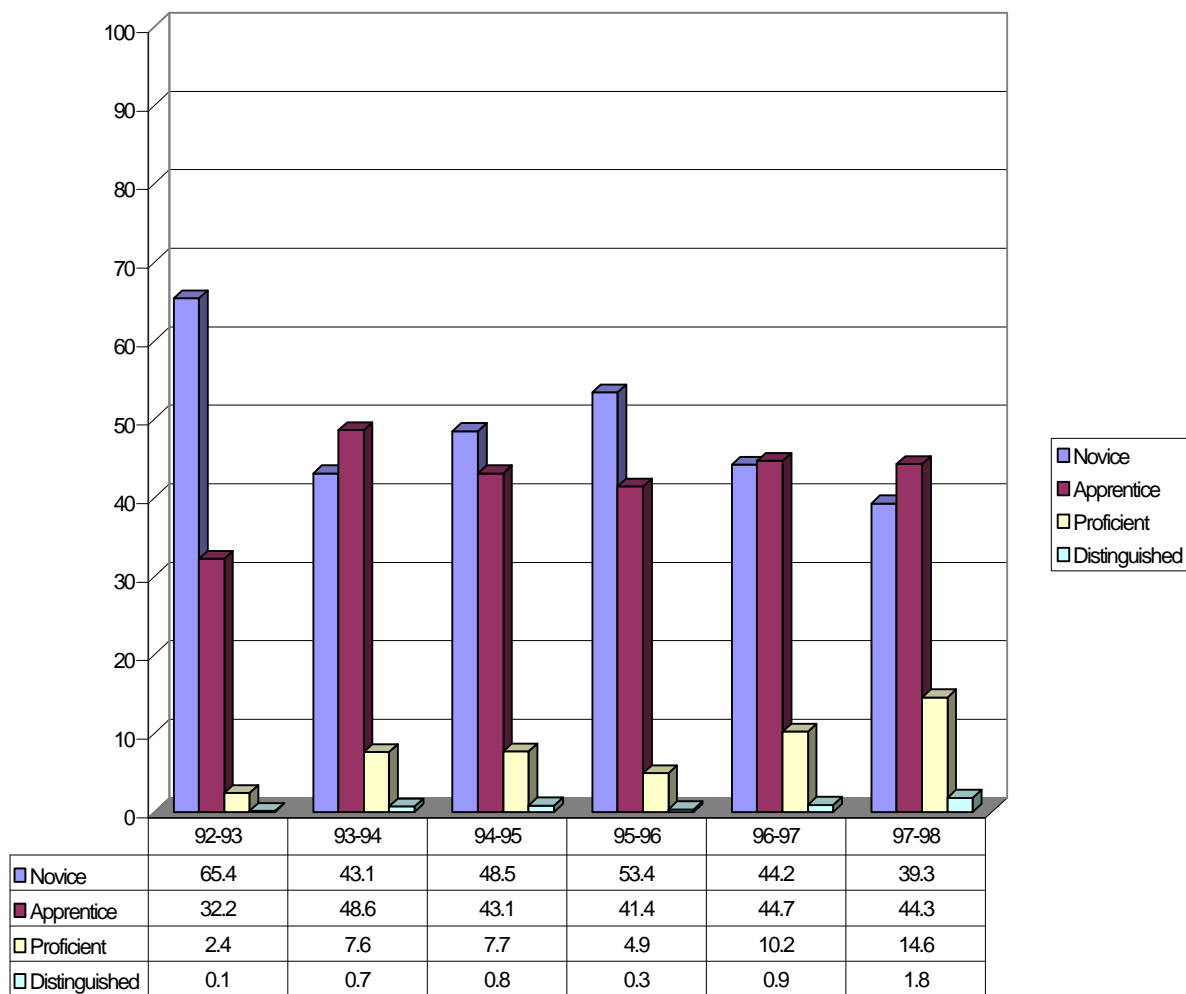


Figure H12. Blacks' Percentage NAPD in Grade 11/12 Social Studies, 1993-1998



Appendix I⁴⁹

KIRIS Cycle 2 and Cycle 3 Subject Area Performance
for Whites by Constructed Response Categories

⁴⁹ Data Tables added to all Figures in Appendix I by KDE.

Figure I1. Whites' Percentage NAPD in Grade 4 Reading, 1993-1998

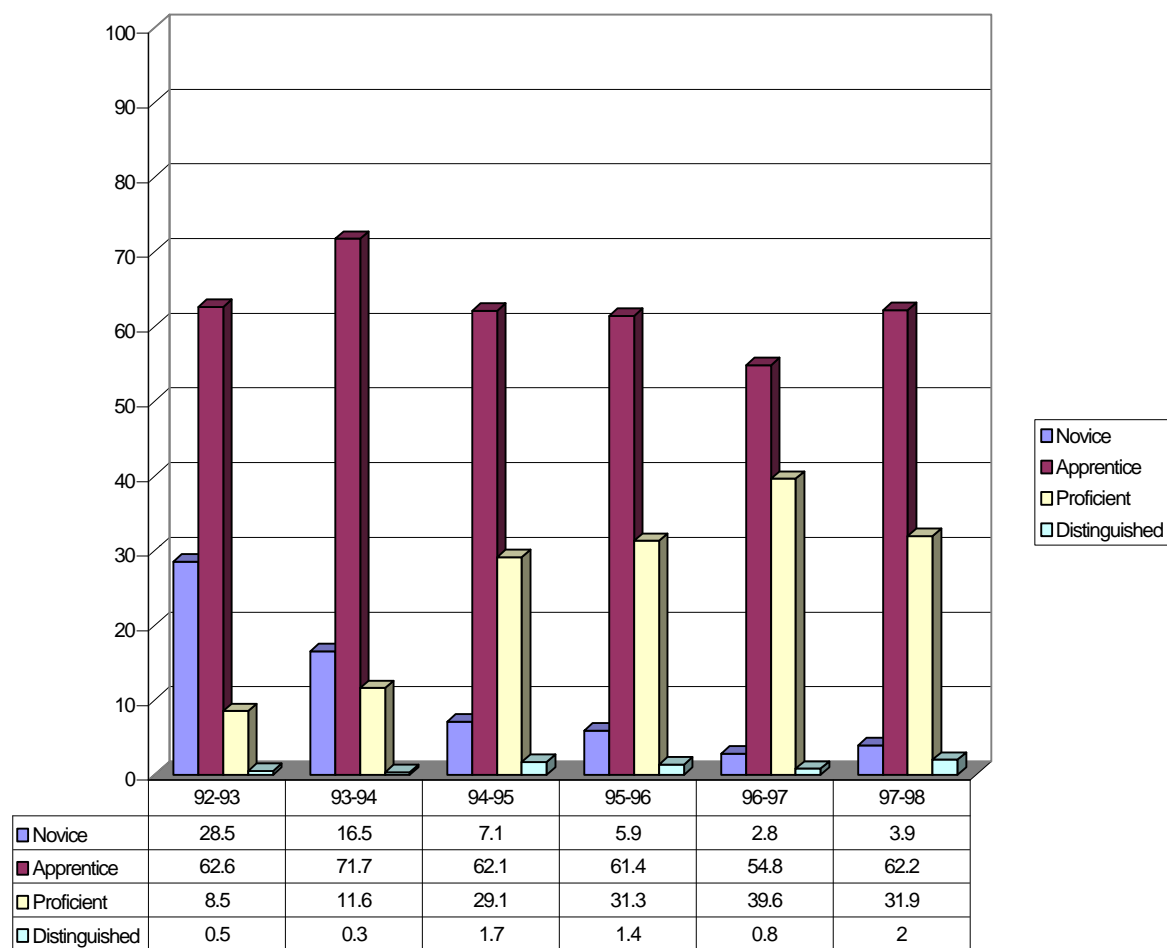


Figure 12. Whites' Percentage NAPD in Grades 4/5 Math, 1993-1998

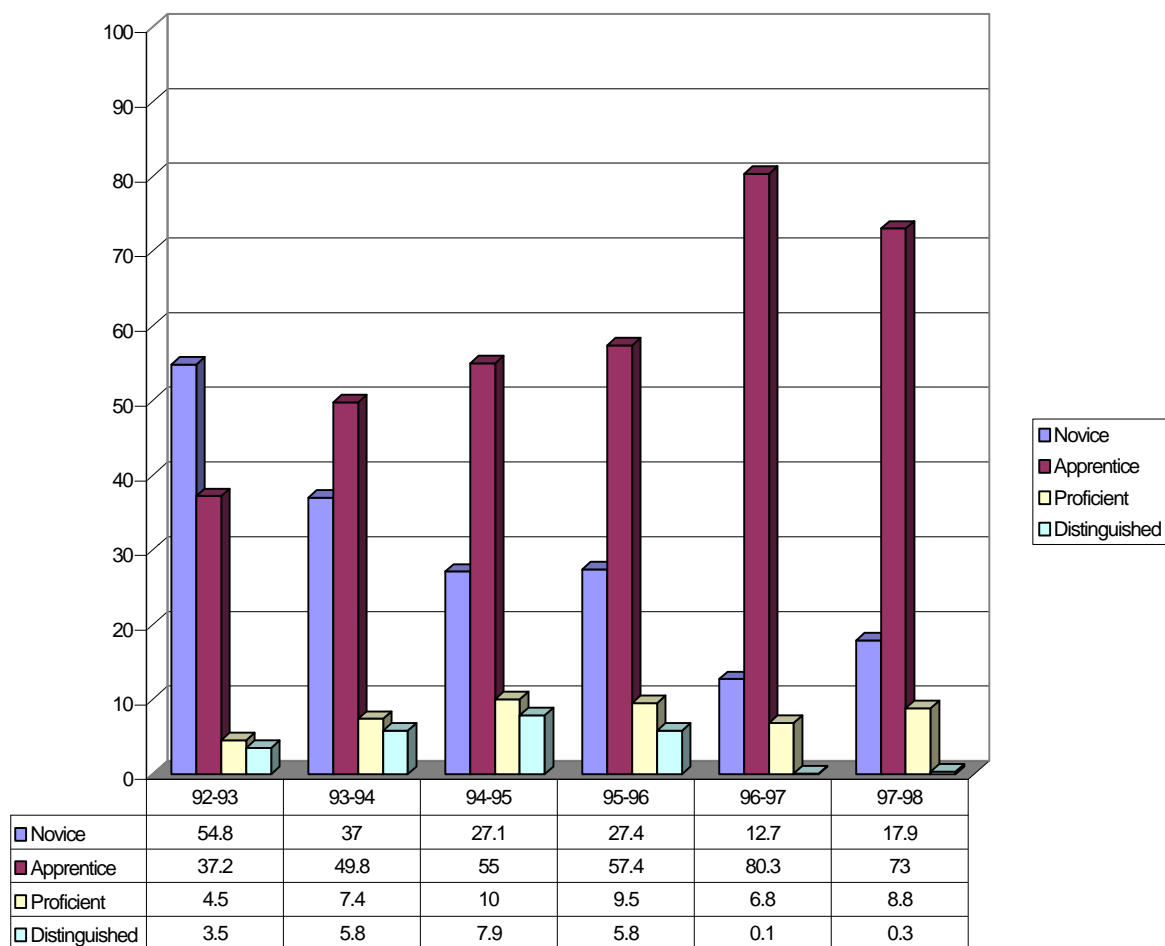


Figure I3. Whites' Percentage NAPD in Grade 4 Science, 1993-1998.

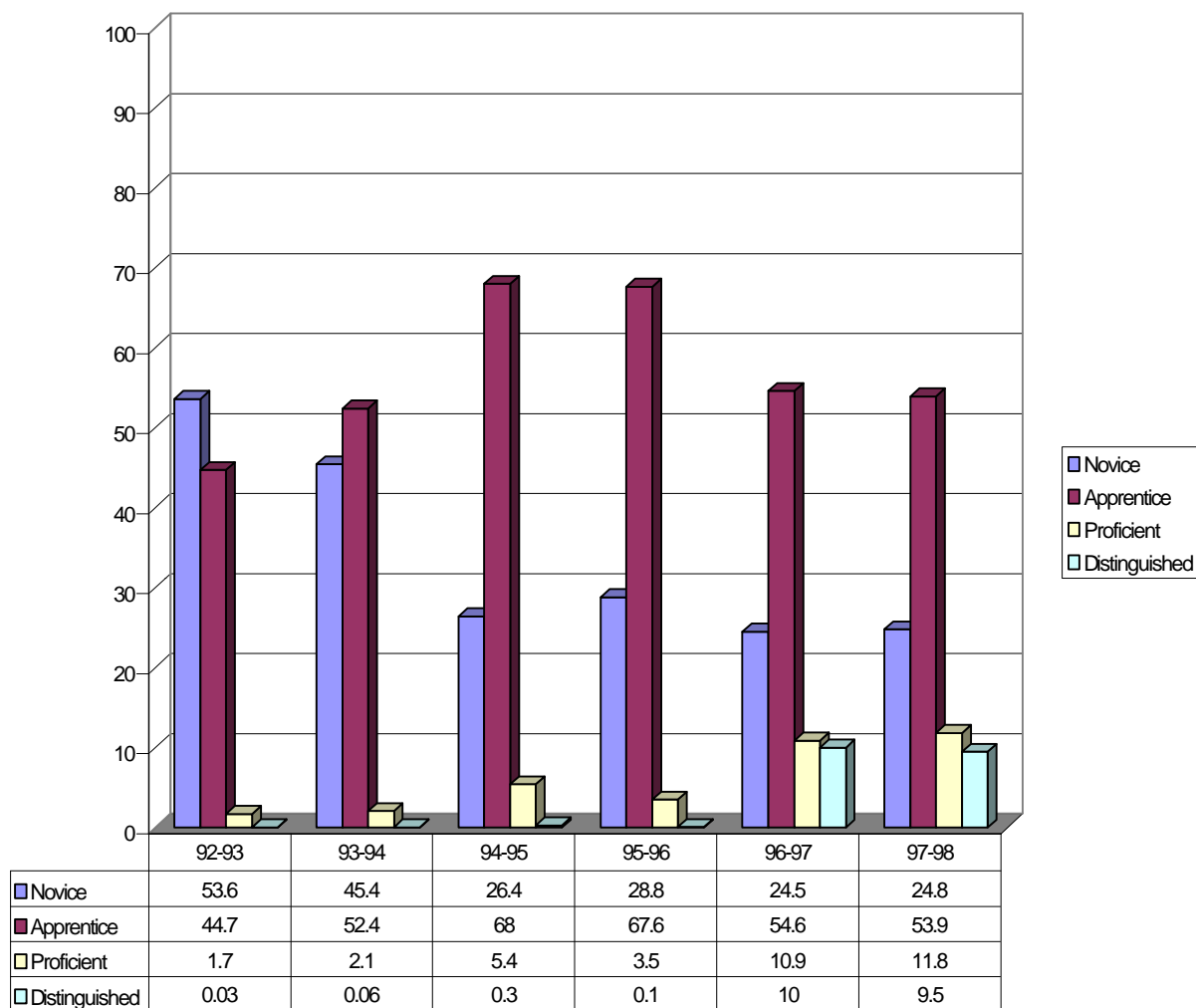


Figure I4. Whites' Percentage NAPD in Grade 4/5 Social Studies, 1993-1998

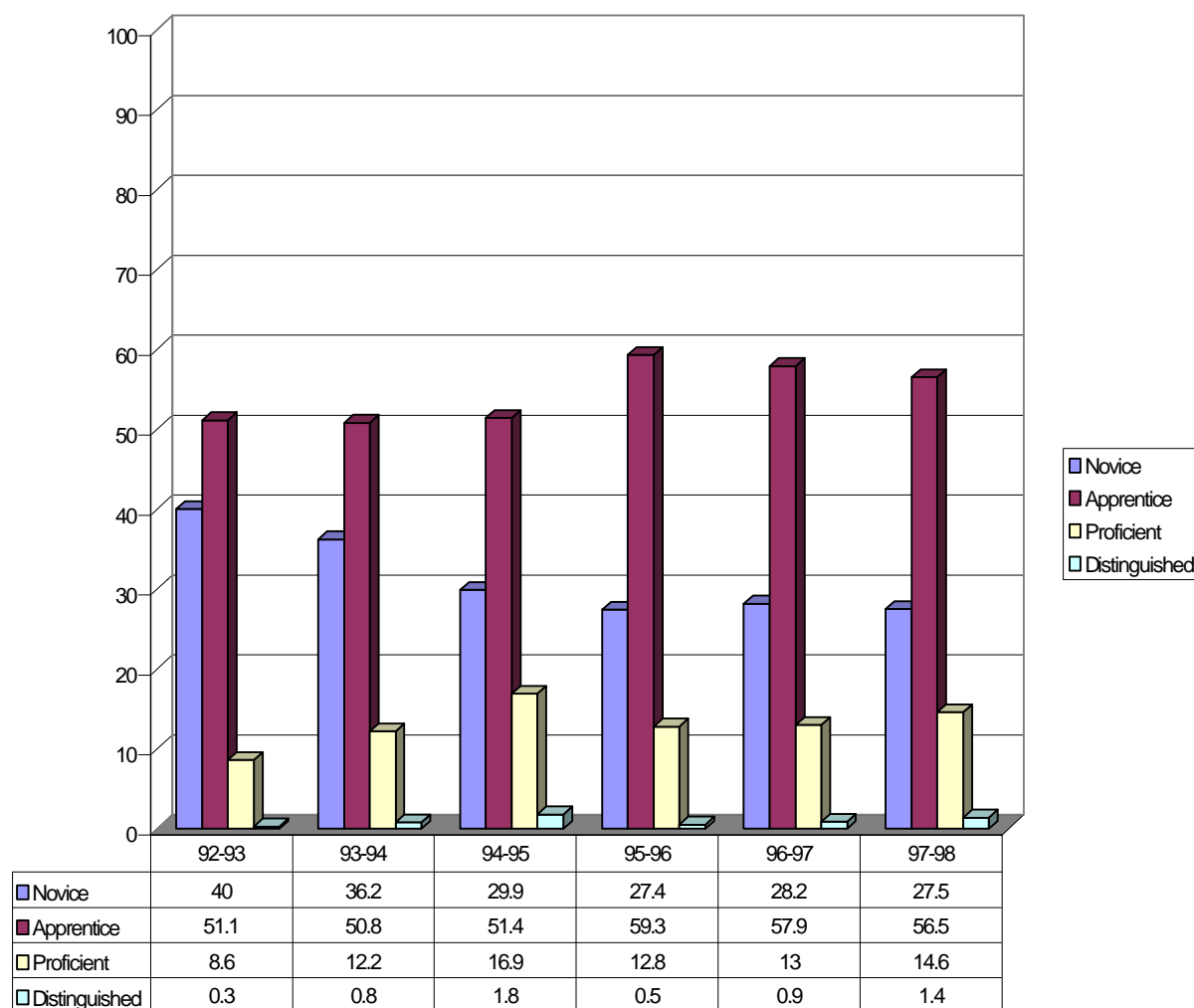


Figure I5. Whites' Percentage NAPD in Grade 7/8 Reading, 1993-1998

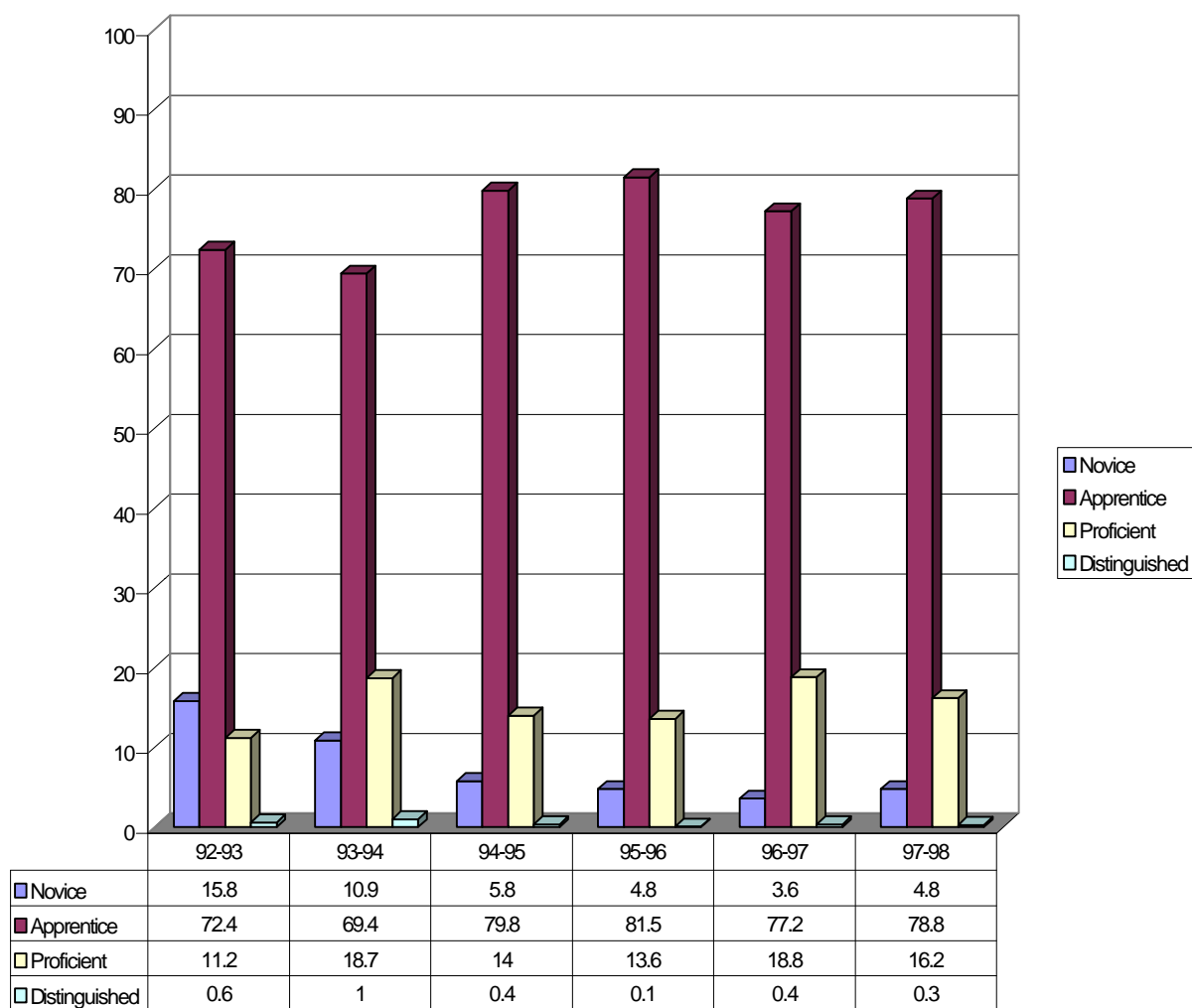


Figure I6. Whites' Percentage NAPD in Grade 8 Math, 1993-1998

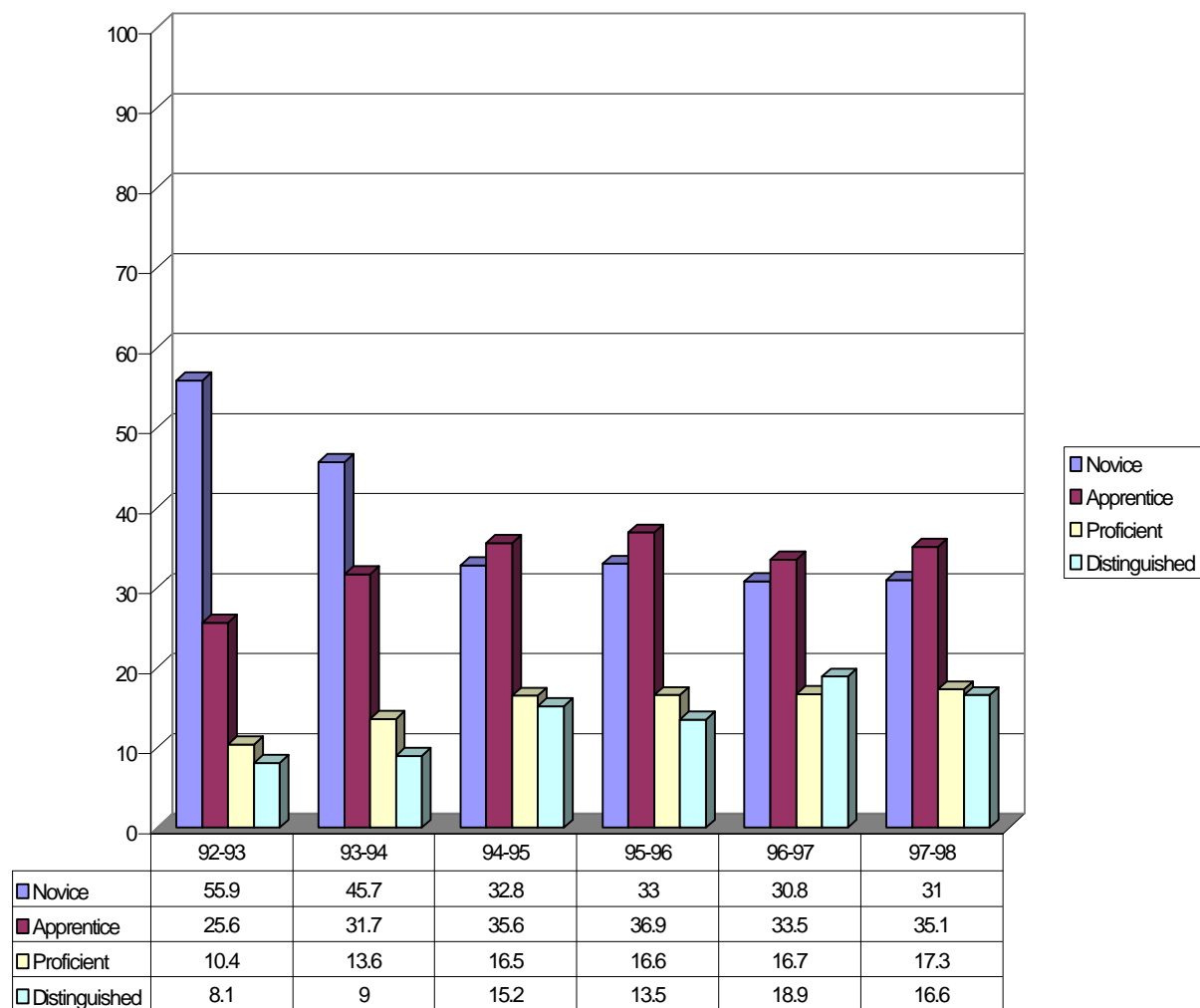


Figure I7. Whites Percentage NAPD in Grade 7/8 Science, 1993-1998

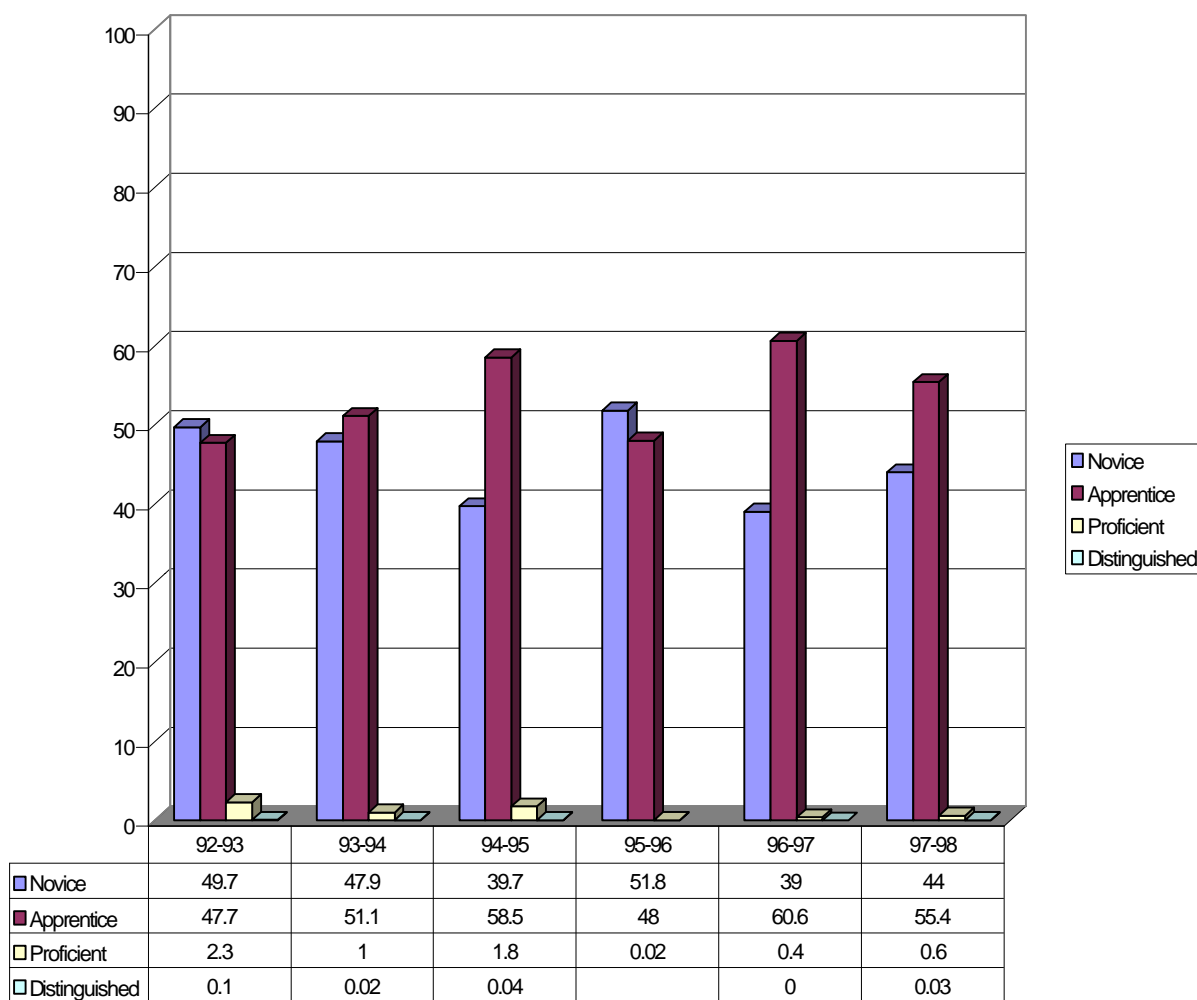


Figure I8. Whites' Percentage NAPD in Grade 8 Social Studies, 1993-1998

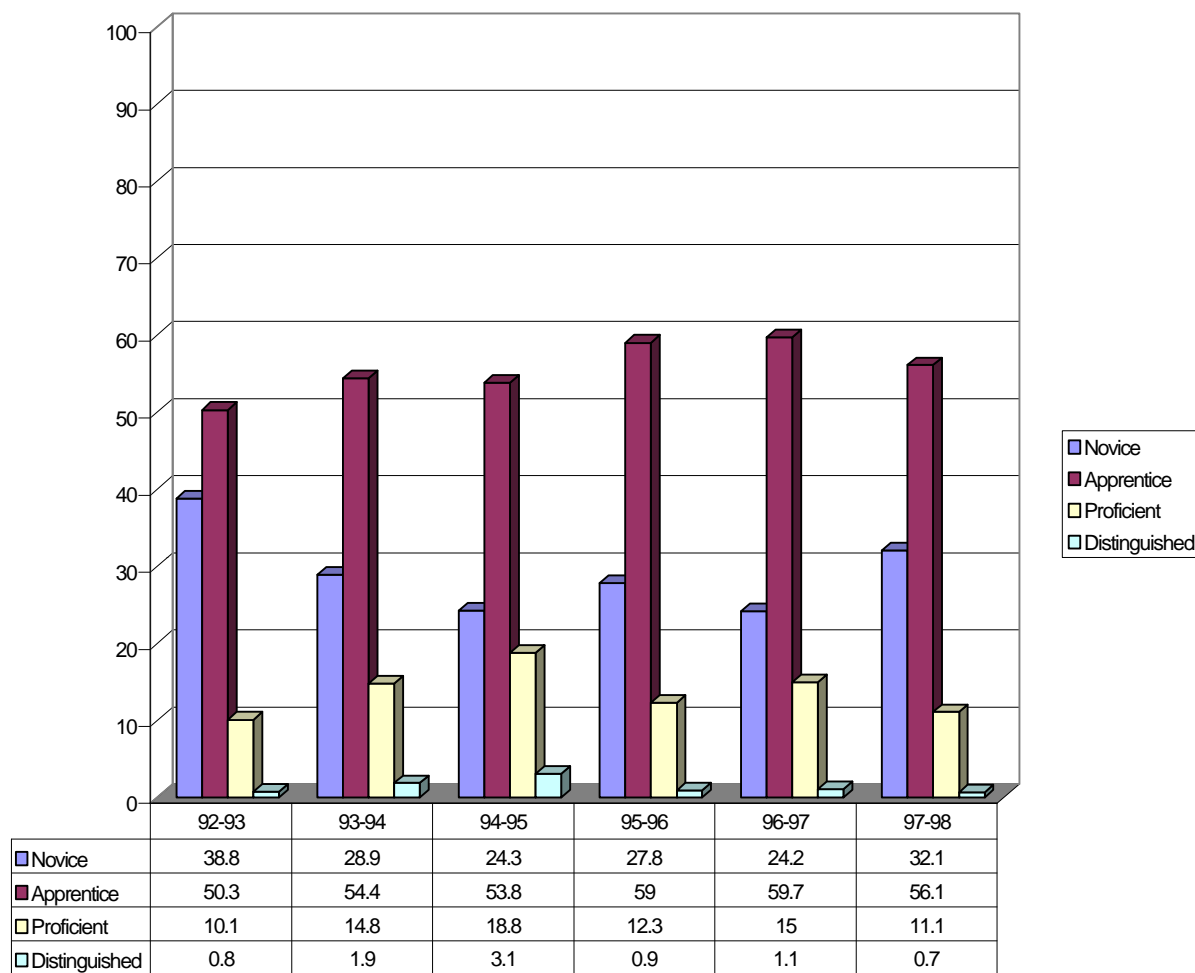


Figure I9. Whites' Percentage NAPD in Grade 11/12 Reading, 1993-1998

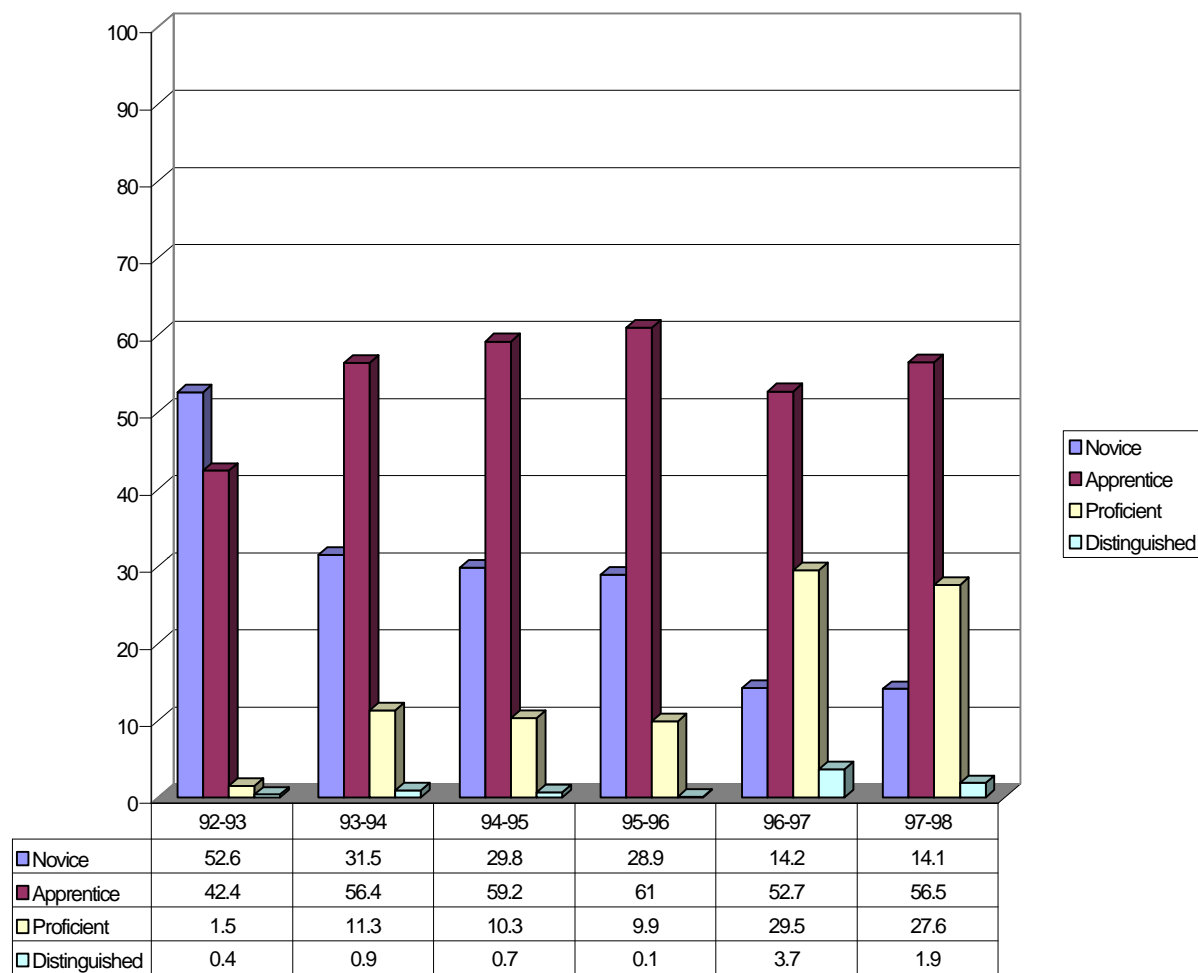


Figure I10. Whites' Percentage NAPD in Grade 11/12 Math, 1993-1998

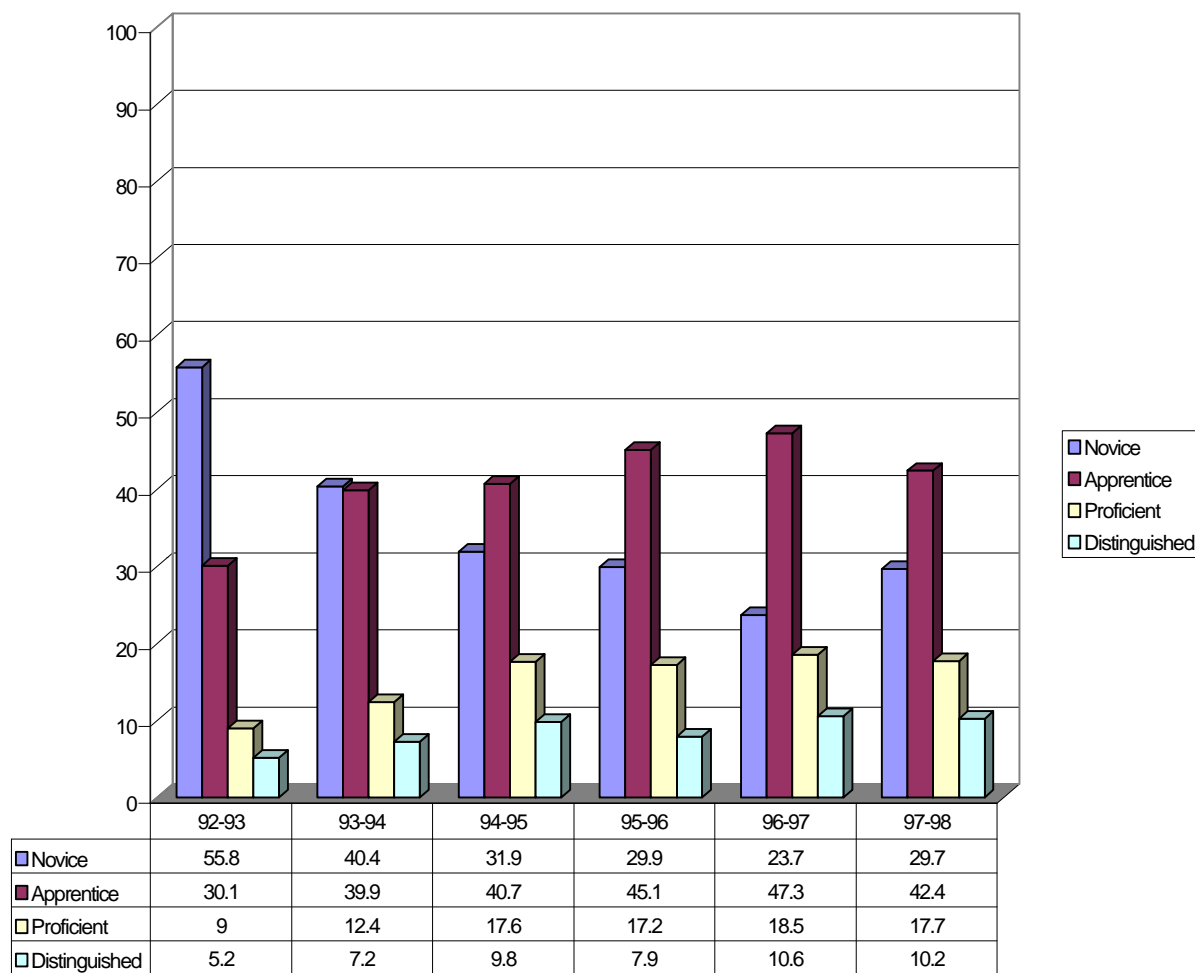


Figure I11. Whites' Percentage NAPD in Grade 11/12 Science, 1993-1998

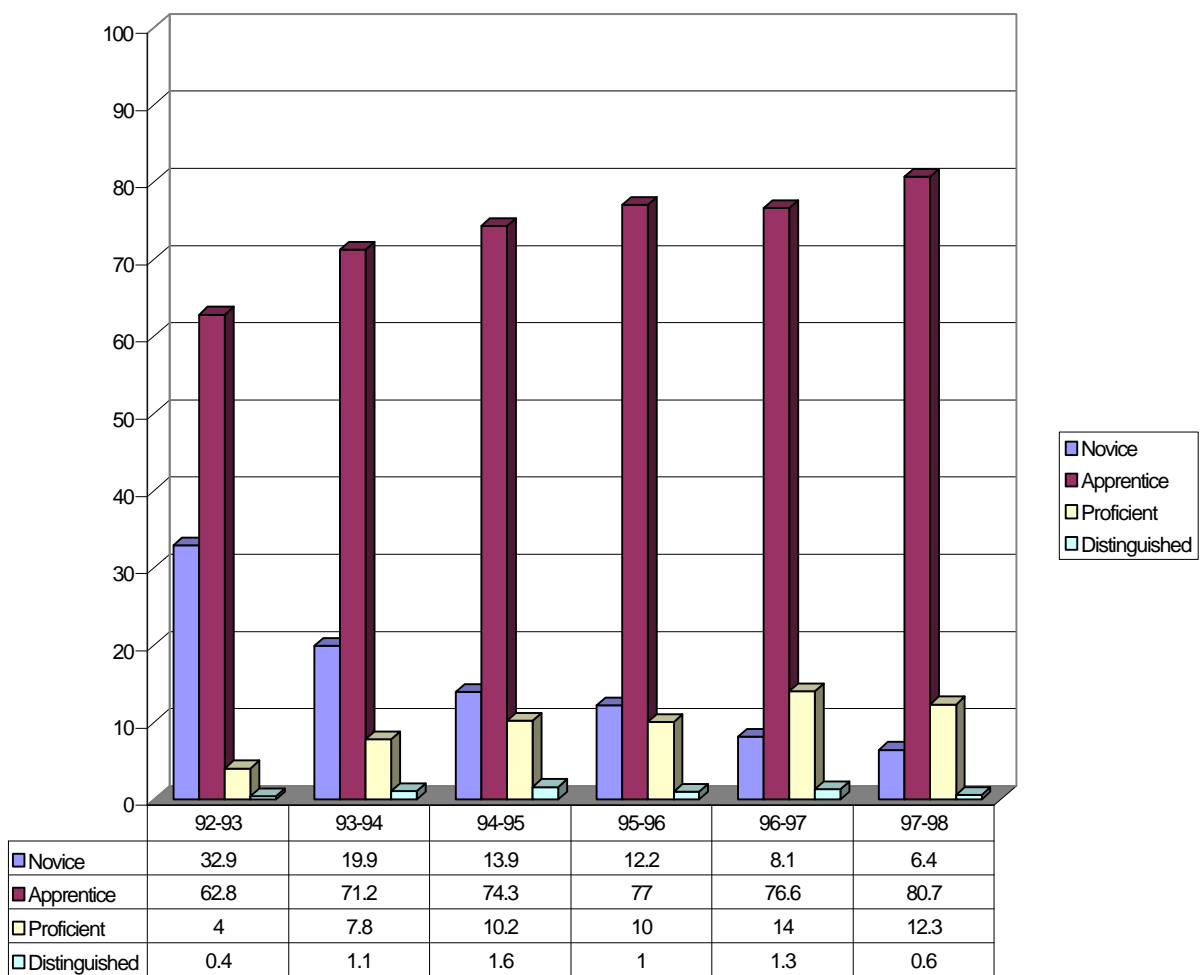
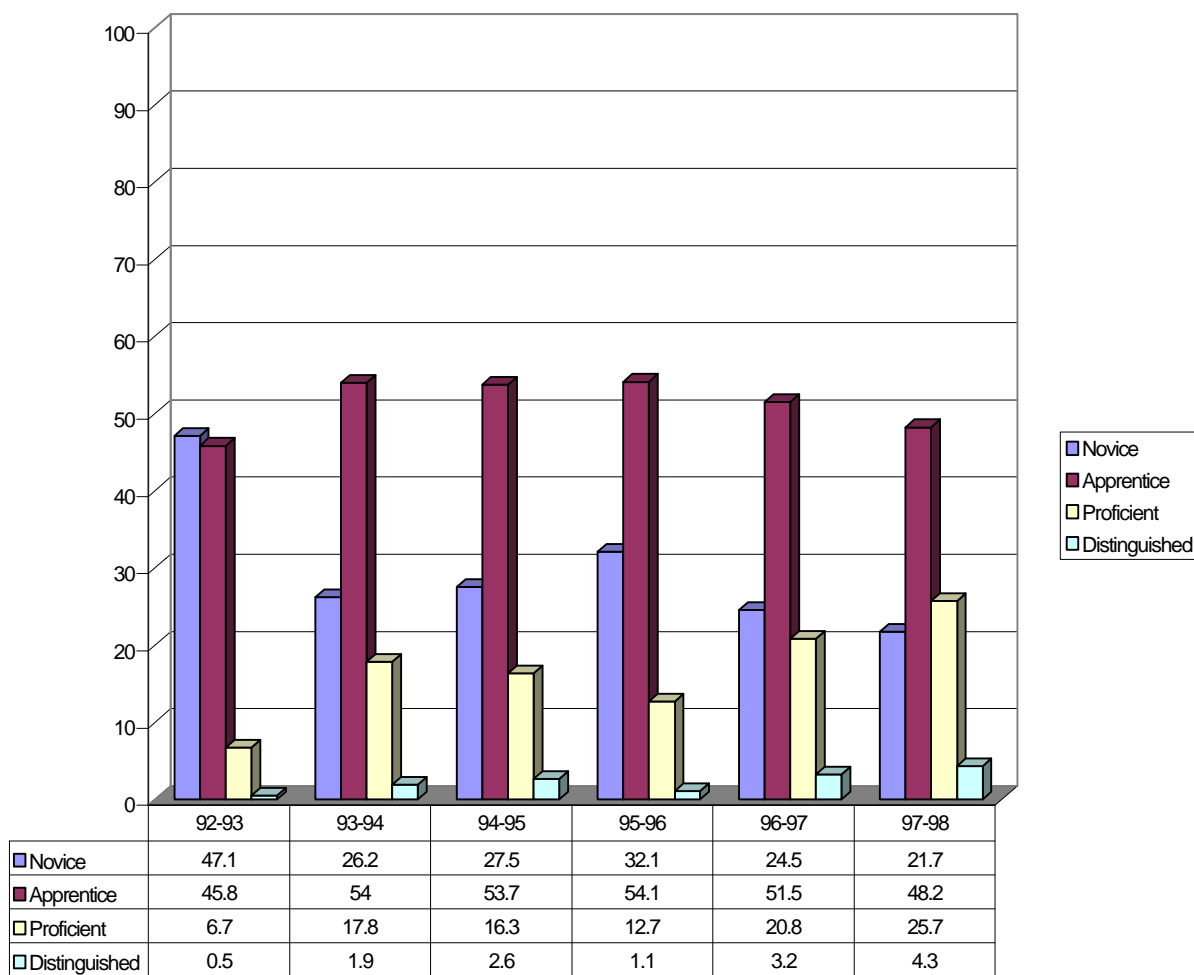


Figure I12. Whites' Percentage NAPD in Grade 11/12 Social Studies, 1993-1998



Appendix J

Outline for Semi-Structured Interview used in Qualitative Study

Semi Structured Interview Questions

Intro

The reason that we wanted to visit your school today is that examination of your students' test scores showed that there was little difference between how your white and African-American students performed on the assessment over the last cycle. This result is especially remarkable since African-American students, on average, perform substantially lower than white students on the assessment, but your students do not and have not over the past four years, when the gap between blacks and whites statewide has increased. What do you think accounts for your success?

Points to make sure and cover:

Curriculum-

Do you have a set curriculum?

How was it agreed upon?

How is information organized?

Do/did issues of race enter into the discussion? How?

Instruction-

Is there a concerted effort at creating instruction practices that are inclusive to all students?

Do instructional practices at your school address the different learning styles? How is this done?

Do you think that students at your school believe that race/ethnicity determines expectations for their performance?

Do you think that race/ethnicity determines how discipline policy is applied?

Professional Development-

What role does professional development play? Has this been a focus?

What's the turnover rate here?

What sorts of things do you do to acclimate a new teacher to this school?

Are issues of race addressed specifically?

Leadership-

How do you get to be more than 20 people pulling different directions?

Have the leaders at this school made racial parity a priority?

Networking-

What sort of connections does this school have with parents? the community? other schools? regional service centers? KDE?

Do you believe that any of these connections have played an important part in your success at racial parity?

Questions added during the interview process on the basis of previous interviews-

What sort of Extended School Services do you offer?

How do students get placed in the program?

What sort of services do the Family and Youth Service Centers offer?

Appendix K

Student Questionnaire Items Used in Regression Models

32. How long have you lived in Kentucky?
1. Less than a year
 2. From one to five years
 3. More than five years
 4. All my life
33. When did you come to this school?
1. This year
 2. Last year
 3. Two years ago
 4. Three or more years ago
34. How many days of school did you miss last month?
1. None
 2. 1 or 2 days
 3. 3 or 4 days
 4. 5 to 10 days
 5. More than 10 days
35. Within the past two years, how many times have you changed schools because you changed where you lived?
1. None
 2. One time
 3. Two times
 4. Three or more times
36. About how many books are there in your home? (Do not count magazines, newspapers, or your schoolbooks.)
1. None or very few (0-10 books)
 2. Enough to fill one shelf (11-25 books)
 3. Enough to fill one bookcase (26-100 books)
 4. Enough to fill two bookcases (101-200 books)
 5. Enough to fill three or more bookcases (more than 200 books)
37. How often do the people in your home speak a language other than English?
1. Never
 2. Sometimes
 3. Always
38. Mark the statement that best describes your overall grades this year.
1. Mostly good or above average
 2. Mostly average
 3. Mostly poor or below average

**ASSESSING RACE AND GENDER SUBGROUP PERFORMANCE
DIFFERENCES IN KIRIS ACCOUNTABILITY CYCLE 2 AND
CYCLE 3 RESULTS**

**A Report for the Kentucky Department of Education
Office of Assessment and Accountability**

by

Douglas Clayton Smith
Western Kentucky University

Dorothy O. Neff
Transylvania University

Jason M. Nemes
Western Kentucky University

TABLE OF CONTENTS

INTRODUCTION	1
Description of KIRIS.....	1
The Gender and Racial Composition of Kentucky Schools	2
Measures of Student Performance.....	2
EXAMINATION OF STUDENT PERFORMANCE BY GENDER AND RACE.....	11
Performance Differences: Are They Significant?	11
Performance Differences: How Big are They?	24
EXPLORATIONS INTO PERFORMANCE DIFFERENCES BY GENDER AND RACE.....	28
<u>Exploratory Study 1: The Geographic Concentration of Gender and Race in Kentucky</u> .	37
The Geographic Concentration of Gender	37
Geographic Distribution of Race	38
Summary of Exploratory Study 1	45
<u>Exploratory Study 2: Qualitative Examination of Schools with Low White/African-</u>	
<u>American Performance Differences</u>	46
School Vignettes.....	48
School A	49
School B	49
School C	49
School D	50
Common Characteristics of School E and F.....	50
School E.....	50
School F.....	50
Additional Characteristics of the School Sample	51
Preliminary Findings of Exploratory Study 2	52
Factors Beyond the Purview of the School Faculty	52
Student Social Class.....	52
Proportion of Black Students in the Schools.....	54
Proportion of Minority Faculty	54
Factors within the purview of the faculty	55
Curriculum	55
Instruction	56
Professional Development	58
Leadership	58
Networking: Parent Involvement, Community Involvement,	
School-to-School Ties	59
Extended School Services and Family Resource Centers.....	60
Summary of Exploratory Study 2.....	60
Exploratory Study 3: Quantitative Examination of Race, Gender and Socioeconomic	
Status.....	61
Summary of Exploratory Study 3.....	77
SUGGESTIONS FOR FUTURE RESEARCH.....	78

<u>Concepts, Variables, Reliability and Validity</u>	78
Individual Level Variables	79
School Level Variables	79
County Level Variables	80
<u>Additional Research Initiatives</u>	81
Qualitative and Quantitative Research.....	81
<u>Qualitative</u>	81
<u>Quantitative</u>	82
Why Triangulate Qualitative and Quantitative Research?.....	82
<u>The Need for Organization</u>	83
REFERENCES	84

APPENDIX A.....	93
APPENDIX B.....	106
APPENDIX C.....	119
APPENDIX D	132
APPENDIX E.....	145
APPENDIX F	158
APPENDIX G	171
APPENDIX H	184
APPENDIX I.....	197
APPENDIX J.....	210
APPENDIX K.....	213